

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 7, 2001, 02:32:15 : Search time 997.91 seconds
(without alignments)
5712.009 Million cell updates/sec

Title: US-09-735-712-1

Perfect score: 603
Sequence: 1 atcgattcaagaccgcacaca.....attgtgaacatgtgtgtga 603

Scoring table: IDENTITY_NTC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_esthum29:*
64: em_esthum30:*
65: em_esthum31:*
66: em_esthum32:*
67: em_esthum33:*
68: em_esthum34:*
69: em_esthum35:*
70: em_esthum36:*
71: em_esthum37:*
72: em_esthum38:*
73: em_esthum39:*
74: em_esthum40:*
75: em_esthum41:*
76: em_esthum42:*
77: em_esthum43:*
78: em_esthum44:*
79: em_esthum45:*
80: em_esthum46:*
81: em_esthum47:*
82: em_esthum48:*
83: em_esthum49:*
84: em_esthum50:*
85: em_esthum51:*
86: em_esthum52:*
87: em_esthum53:*
88: em_esthum54:*
89: em_esthum55:*
90: em_esthum56:*
91: em_esthum57:*
92: em_esthum58:*
93: em_esthum59:*
94: em_esthum60:*
95: em_esthum61:*
96: em_esthum62:*
97: em_esthum63:*
98: em_esthum64:*
99: em_esthum65:*
100: em_esthum66:*
101: em_esthum67:*
102: em_esthum68:*
103: em_esthum69:*
104: em_esthum70:*
105: em_esthum71:*
106: em_esthum72:*
107: em_esthum73:*
108: em_esthum74:*
109: em_esthum75:*
110: em_esthum76:*
111: em_esthum77:*
112: em_esthum78:*
113: em_esthum79:*
114: em_esthum80:*
115: em_esthum81:*
116: em_esthum82:*

```
117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*
```

```
190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	466.2	77.3	508	16	AI149899 g443n06.x
C 2	450.6	74.7	516	7	AA436088 zud03a08.r
C 3	347.7	57.5	387	7	AA416972 zc94h05.s
C 4	340.2	56.4	382	7	AA435988 zud03a08.s
C 5	291.8	48.4	394	11	AA758635 ab67b04.s
C 6	283.8	47.1	484	163	BE107659 U1-R-BT1-
C 7	267.4	44.3	410	7	AA411806 zc67a03.s
C 8	245.4	40.7	389	11	AA781801 a158q01.s
C 9	242.8	40.3	415	7	AA470059 zc94h05.r
C 10	207	34.3	482	14	AI002083 ct38b02.s
C 11	141	23.4	332	147	BE319786 uy63h11.x
C 12	131.8	21.9	601	138	BE638317 EST00003
C 13	126.2	20.9	224	138	BE638325 EST00022
C 14	115.6	19.2	538	224	AAQ108532 CIT-HSP-2
C 15	86.4	13.9	1071	106	AL544561 ah41a12.s
C 16	86.4	13.9	1071	106	AL544561 ah41a12.s
C 17	83	13.8	793	154	BE6484817 602505511
C 18	83	13.8	218	138	BE638321 EST00013
C 19	82.2	13.6	704	153	BE571626 602592934
C 20	80.4	13.3	508	4	AA234138 zc51b06.r
C 21	80.4	13.3	562	7	AA418443 zc92e05.r
C 22	70.6	11.7	421	257	B86842 RPII1-26F1
C 23	61.2	10.1	887	106	AL531049 AL531049
C 24	59.4	9.9	715	155	BE538851 602568296
C 25	55.6	9.2	579	136	BE513276 601315340
C 26	55.6	9.2	947	192	AK008652 Mus muscu
C 27	55.4	9.2	470	102	BE607884 wf43q11.x
C 28	55	9.1	528	138	BE675149 7f03d02.x
C 29	55	9.1	904	172	BE602463 602275469
C 30	54.6	9.1	687	143	BE656859 7k10g10.x
C 31	54.4	9.0	1504	192	AK003110 Mus muscu
C 32	53	8.8	1184	192	AK017928 Mus muscu
C 33	52.6	8.7	515	153	BE643437 602506331
C 34	52.6	8.7	594	32	AV717594 AV717594
C 35	52.6	8.7	597	32	AV716310 AV716310
C 36	52.6	8.7	619	32	AV661743 AV661743
C 37	52.6	8.7	800	155	BE546745 602574115
C 38	52.6	8.7	735	32	AV715678 AV715678
C 39	52.4	8.7	504	150	BE522968 U1-R-C2P-
C 40	52.4	8.7	755	104	AI950360 wp10b06.x
C 41	52.2	8.7	736	137	BE563951 601348160
C 42	52.2	8.7	860	106	AL551751 AL551751
C 43	52	8.6	534	151	BE591089 7h52g02.x
C 44	51.4	8.5	559	19	AI391038 mc10b04.Y
C 45	50.8	8.4	482	151	BE600732 265501.MA

ALIGNMENTS

RESULT 1	508 bp	EST	10-NOV-1998
AI149899	508 bp	EST	10-NOV-1998
qf43n06.x1 Soares, testis_NHT Homo sapiens cDNA clone IMAGE:1752827			
3', similar to SW:CD20_HUMAN P11836 B-LYMPHOCTE ANTIGEN CD20 ;			
LOCUS			
DEFINITION			
AI149899			
AI149899			
AI149899.1	GI:3678368		
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
human			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
REFERENCE			
1 (bases 1 to 508)			
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.			
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),			
Tumor Gene Index			

JOURNAL COMMENT

Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: ccgaps-remail.nih.gov
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CCGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: www-bio.llnl.gov/bdrr/image/image.html
 Insert length: 742 Std Error: 0.00
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 455.
 Location/Qualifiers
 1. 508

FEATURES

source
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1752827"
 /clone_lib="Soares, testis_NHT"
 /sex="male"
 /lab_host="DH10B"
 /note="vector: pT7AD-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dt) primer [5', TCTTACCAATCGAAGTGGAGCGGCCCAATTTTCTTTTCTT 3']
 Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia) and digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization to Col5, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT

186 a 91 c 89 g 142 t

ORIGIN

Query Match 77.3%; Score 466.2; DB 16; Length 508;
 Best Local Similarity 99.4%; Pred. No. 5e-109;
 Matches 468; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY	133	agaaaaaagaaatccttagagacatcgaatcctgtttggaattatgacctttctt	192
DB	508	AGAAAAATGAAATCTTAGGACATCCAGATCCTGTTGAAATATGACCTTTCTTTT	449
QY	193	ggaattatctcttcccttccctgtttaaaccatcgaagttcccttataattctt	252
DB	448	GGAATTATCTCTCTTTCACCTTTGTTAAACATATCCAGAGTTCCCTTATTTCTT	389
QY	253	tcaagatccatctcgtgggtcgtgtttgttcaattatccttgagccttcaattgca	312
DB	388	TCAAGATATCCATCTGAGGCTGTGTTTGTTCATTAATCTGAGCCTTCAATTGCA	329
QY	313	gtgaaaaagaaacacacaaagaaactcgataatattgagccgaataaagaattctt	372
DB	328	GTGAAAAAGAAACACACAAAGAACTGTATATATGAGCCCAATTAATTAATTTCTT	269
QY	373	gacctgagacaaatagctggaatcattctcctacatttggttccatccatgacaaac	432
DB	268	GCCTGGAGCAATAGCTGGAATCATTTCTCTCACTTTGTTTCATCTGATCAAAAC	209
QY	433	tacatttggttattctcacaacaaatagtcagtgtgaagcgttactgtcttcttg	492
DB	208	TACATTGTGTGTTATCTCTCAACAAATAGTCAGTAAAGCTGTACTGTCCTCTTG	149
QY	493	ggaatttgattatcatgagactttagaatattgaattatcatctctctcttc	552
DB	148	GGAATTTTGATTAATGATGATGACTTTCAGCATTAATTAATTAATTTCTTCTTC	89
QY	553	tcaatttggtggtccactcagagagattgattgttaaacattgttga 603	
DB	88	TCAATTTTGGGGTCCCTCAGAGAGATTGATTGTGAACAATGTTTGA 38	

RESULT 2

LOCUS	AA411806	410 bp	mRNA	EST	12-NOV-1997
DEFINITION	z67a03.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:273723', similar to SW:CD20_HUMAN P11836 B-LYMPHOCYTE ANTIGEN CD20 ;, mRNA sequence.				
ACCESSION	AA411806				
VERSION	AA411806.1 GI:2070377				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.				
AUTHORS	Mammalia; (bases 1 to 410) Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marris, M., Martin, V., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R. WashU-Merck EST Project 1997 Unpublished (1997) Contact: Wilson RK Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@wustl.edu This clone is available royalty-free through LIND ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Insert Length: 608 Std Error: 0.00 Seq primer: -41ml3 fwd. Ef from Amersham High quality sequence stop: 401. Location/Qualifiers 1..410 /organism="Homo sapiens" /db_xref="GDB:5924284" /db_xref="taxon:9606" /clone="IMAGE:727372" /clone_lib="Soares_testis_NHT" /sex="male" /lab_host="DH10B" /note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGGCCGCCCAATTTTATTTTATTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo."				
BASE COUNT	147 a 68 c 76 g 119 t				
ORIGIN					
Query Match	44.3%; Score 267.4; DB 7; Length 410;				
Best Local Similarity	96.1%; Pred. No. 3.6e-58;				
Matches 274; Conservative	0; Mismatches 11; Indels 0; Gaps 0;				
OY	81	acttcagccacgaccttccaactcaacgaagcccttgcaaaatatttgcgtaga	140		
Db	410	ACTTCAGCCACGACCTTTCACATCAACAAAGCCCTTCGCAAAATATTTGCTAG	351		
OY	141	gaaactcttaggaatcccaagatccctgtttggaattatgaccttctctttgag	200		
Db	350	GAAAACTTTAGGAGATATCCAGATCCCTCTTTGGAAATATGACCTTTCTTTGG	291		
OY	201	cttccttccacctgtcttaaaacacacacaaaggttcctcttatattcttcag	260		
Db	290	CTTCCTTTTCACCTTGTTAAACCAATCCAAAGTTCCCTTTATATTTCTTCA	231		
OY	261	tcacattctggggctctgtttgttcattaatctcgtgagacctcctaattgcag	320		
Db	230	TCCATTTGGGGCTCTGTTTGTTCATTAATCTTGAGACCTTCCCAATTTGCA	171		
OY	321	aaaaaacacagaaactcgtgataataattgagccgaataatgatctc	365		

Db	170	AAAAACGACAGAACTCTGGCAATTTTGTATTCATTCATGACTTT	126
RESULT	8		
LOCUS	AA781801/c		
DEFINITION	AA781801	389 bp	mRNA
ACCESSION	AA781801		EST
VERSION	AA781801.1	GI:2841132	31-DEC-1998
KEYWORDS	EST		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 389)		
JOURNAL	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
COMMENT	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),		
	Tumor Gene Index		
	Unpublished (1997)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgaps-remail.nih.gov		
	cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo		
	, Ph.D.		
	cDNA Library Arrayed by: Greg Lennon, Ph.D.		
	DNA sequencing by: Washington University Genome Sequencing Center		
	Clone distribution: NCI-CGAP clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	www-bio.llnl.gov/dbp/image/image.html		
	Insert Length: 983 Std Error: 0.00		
	Seq primer: -40m13 fwd. ET from Amersham		
	High quality sequence stop: 311.		
FEATURES	Location/Qualifiers		
SOURCE	1..389		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="1375056"		
	/clone_1bp="Soares_tests_NHT"		
	/sex="male"		
	/lab_host="DH10B"		
	/note="Vector: pT73D-Pac (Pharmacia) with a modified		
	polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA		
	was prepared from mRNA obtained from Clontech Laboratories		
	, Inc., and primed with a Not I - oligo(dT) primer [5'		
	TGTATCCAACTCGAAGGAGGAGCGGCCGCCAATTTTTTTTTTTT 3'].		
	Double-stranded cDNA was ligated to Eco RI adaptors		
	(Pharmacia), digested with Not I and cloned into the Not I		
	and Eco RI sites of the modified pT73 vector. Library		
	went through one round of normalization to Cot5, and was		
	constructed by Bento Soares and M. Fatima Bonaldo."		
BASE COUNT	140 a	66 c	68 g
ORIGIN			115 t
Query Match	40.7%;	Score 245.4;	DB 11; Length 389;
Best Local Similarity	95.8%;	Pred. No. 1.5e-52;	
Matches	252;	Conservative 0;	Mismatches 11; Indels 0; Gaps 0;
Db	103	actcaagcccttcgcaaaatattctgctgtagaaaatgaanaactcttaggagataccag	162
Db	389	ACTCAAAACCCCTTGCAAAAATATTTCCTGTGAAAAATGAAAAATCTTAGGACATATCAG	330
Db	163	atccggttgaatatgaacttcttccttggagttatctccttcaactgtgtaaa	222
Db	329	ATCCGTTTGGAAATATACCTTTCTTTTGGAGTATATCTTTCACCTTGTAAAA	270
Db	223	ccatatccaaggttccttataatcttcacagatatccatctcgtggtctgtttg	282
Db	269	CCATATCCAAAGTTTCCCTTATATTCTTTCAGAGATATCATTCGTGGCGCTGTATTG	210
Db	283	ttaataattcttgagacctctctaattgcagtgaaagaanaaccacagaactctgata	342

DB	209	TTTCATTAATTCGTGAGACCTTCCTCAATTCGAGTGAAGAAAAAGAACACAGAACTCGGGA	150
QY	343	atattgagccgaataatgaatct 365	
DB	149	ATTGTGATTACATTCGATGACTTT 127	
RESULT	9		
LOCUS	AA470059	415 bp	EST
DEFINITION	z194h05.r1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:730041		09-NOV-1997
ACCESSION	AA470059		
VERSION	AA470059.1	GI:2197368	
KEYWORDS	EST.		
ORIGIN	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
	1 (bases 1 to 415)		
	Hillier, L., Allen, M., Bowles, L., Dubnue, T., Geisel, G., Jost, S.,		
	Krivan, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin		
	, J., Moore, B., Schellenberg, K., Stepec, M., Tan, F., Theising, B.,		
	White, Y., Wyllie, T., Waterston, R. and Wilson, R.		
TITLE	WashU-NCI human EST project		
JOURNAL	Unpublished (1997)		
COMMENT	Contact: Wilson RK		
	Washington University School of Medicine		
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108		
	Tel: 314 286 1800		
	Fax: 314 286 1810		
	Email: est@watson.wustl.edu		
	This clone is available royalty-free through LINDL ; contact the		
	IMAGE Consortium (info@image.lindl.gov) for further information.		
	Insert length: 656 Std Error: 0.00		
	Seq primer: -28m13 rev2 ET from Amersham		
	High quality sequence stop: 353.		
FEATURES	Location/Qualifiers		
SOURCE	1. 415		
	/organism="Homo sapiens"		
	/db_xref="GDB:5926570"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:730041"		
	/clone_lib="Soares_testis_NHT"		
	/sex="male"		
	/lab_host="DH10B"		
	/note="vector: pT7T3D-Pac (Pharmacia) with a modified		
	polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA		
	was prepared from mRNA obtained from Clontech Laboratories		
	, Inc., and primed with a Not I - oligo(dT) primer [5'		
	TGTTACCAATCTGAAGTGGAGCGCGCCCAATTTTTTTTTTTT 3'].		
	Double-stranded cDNA was ligated to Eco RI adaptors		
	(Pharmacia), digested with Not I and cloned into the Not I		
	and Eco RI sites of the modified pT7T3 vector. Library		
	went through one round of normalization to Cot5, and was		
	constructed by Bento Soares and M. Fatima Bonaldo. "		
BASE COUNT	129 a 96 c 67 g 123 t		
ORIGIN			
Query Match	40.3%; Score 242.8; DB 7; Length 415;		
Best Local Similarity	77.3%; Pred. No. 7,1e-52;		
Matches	357; Conservative 0; Mismatches 2; Indels 103; Gaps 1;		
QY	1	atggaattcaagcaccgcacacagtcggtglttctgtgatttcctcagaatacactgct 60	
DB	57	ATGGATTCAAGCACCGCACACAGTCCGGTGTTTCGTATTTCCACGAAATCACTGCT 116	
QY	61	tcaagaatgagttcaagaagaactttcagccaagacctttcaatcacaagcccttgaa 120	
DB	117	TCAGAAATGAGTCCACAGAACTTTCAGCGCACGACTTTTCAACTCAAGCCCTTGGCA 176	

DB	177	181	209	241	209	301	254	361	314	421	374	RESULT
121	177	181	209	241	209	301	254	361	314	421	374	10
aaatttgcagagaagaatgaaatcttagagactacagatccctgttggaattatg	AAATTATTGTCAGAGAAAATGAAATCTTAGG	accttctcttgtagagtaiccttcccttccacttgtaaaacacataccaagggttccc	tttatattcttcttagagatalccatctctcggggctctgtttgttcatatattcggagcc	-----GGATATTCATTCTGGGGGCTCTGTTTGTTCATTAAATTCGGAGACC	ttcccaattgcagtagaagaagaacccacagaactctgataatatttagccggaataatg	TTCCCAATTTGCGAGTAAAAAGAAAAACCCAGAAACTCTGATATATTTTtagccggaataatg	aatcttcttagtgcctcttagagacaatagctggaatcattcttccctccacatttggttcac	aatttcttagtgcctcttagagacaatagctggaatcattcttccctccacatttggttcac	AATTTCTTAGTGCCCTGGAGCAATAGCTGGAATTCATTCCTTCACATTTGTTTCATC	ctagatcaaaactacatttgtgttatcttccacaaatagc	CTAGATCAAAACTACATTGTGTGTTATTCACCAAAATAGCT	482 bp mRNA EST 27-AUG-1998
LOCUS	AI002083	AI002083	AI002083	AI002083	AI002083	AI002083	AI002083	AI002083	AI002083	AI002083	AI002083	AI002083
DEFINITION	LOCUS	LOCUS	LOCUS	LOCUS	LOCUS	LOCUS	LOCUS	LOCUS	LOCUS	LOCUS	LOCUS	LOCUS
3', mRNA sequence.	3', mRNA sequence.	3', mRNA sequence.	3', mRNA sequence.	3', mRNA sequence.	3', mRNA sequence.	3', mRNA sequence.	3', mRNA sequence.	3', mRNA sequence.	3', mRNA sequence.	3', mRNA sequence.	3', mRNA sequence.	3', mRNA sequence.
ACCESSION	AI002083	AI002083	AI002083	AI002083	AI002083	AI002083	AI002083	AI002083	AI002083	AI002083	AI002083	AI002083
VERSION	AI002083.1	AI002083.1	AI002083.1	AI002083.1	AI002083.1	AI002083.1	AI002083.1	AI002083.1	AI002083.1	AI002083.1	AI002083.1	AI002083.1
KEYWORDS	EST.	EST.	EST.	EST.	EST.	EST.	EST.	EST.	EST.	EST.	EST.	EST.
SOURCE	human.	human.	human.	human.	human.	human.	human.	human.	human.	human.	human.	human.
ORGANISM	Homo sapiens	Homo sapiens	Homo sapiens	Homo sapiens	Homo sapiens	Homo sapiens	Homo sapiens	Homo sapiens	Homo sapiens	Homo sapiens	Homo sapiens	Homo sapiens
REFERENCE	Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	Enkaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT	Tumor Gene Index	Tumor Gene Index	Tumor Gene Index	Tumor Gene Index	Tumor Gene Index	Tumor Gene Index	Tumor Gene Index	Tumor Gene Index	Tumor Gene Index	Tumor Gene Index	Tumor Gene Index	Tumor Gene Index
	Unpublished (1997)	Unpublished (1997)	Unpublished (1997)	Unpublished (1997)	Unpublished (1997)	Unpublished (1997)	Unpublished (1997)	Unpublished (1997)	Unpublished (1997)	Unpublished (1997)	Unpublished (1997)	Unpublished (1997)
	Contact: Robert Strausberg, Ph.D.	Contact: Robert Strausberg, Ph.D.	Contact: Robert Strausberg, Ph.D.	Contact: Robert Strausberg, Ph.D.	Contact: Robert Strausberg, Ph.D.	Contact: Robert Strausberg, Ph.D.	Contact: Robert Strausberg, Ph.D.	Contact: Robert Strausberg, Ph.D.	Contact: Robert Strausberg, Ph.D.	Contact: Robert Strausberg, Ph.D.	Contact: Robert Strausberg, Ph.D.	Contact: Robert Strausberg, Ph.D.
	Email: cgapbs-remail.nih.gov	Email: cgapbs-remail.nih.gov	Email: cgapbs-remail.nih.gov	Email: cgapbs-remail.nih.gov	Email: cgapbs-remail.nih.gov	Email: cgapbs-remail.nih.gov	Email: cgapbs-remail.nih.gov	Email: cgapbs-remail.nih.gov	Email: cgapbs-remail.nih.gov	Email: cgapbs-remail.nih.gov	Email: cgapbs-remail.nih.gov	Email: cgapbs-remail.nih.gov
	cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo	cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo	cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo	cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo	cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo	cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo	cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo	cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo	cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo	cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo	cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo	cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
	, Ph.D.	, Ph.D.	, Ph.D.	, Ph.D.	, Ph.D.	, Ph.D.	, Ph.D.	, Ph.D.	, Ph.D.	, Ph.D.	, Ph.D.	, Ph.D.
	cDNA Library Arrayed by: Greg Lennon, Ph.D.	cDNA Library										

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

GenCore version 4.5
Copyright (c) 1993 - 2000, Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 7, 2001, 02:33:15; Search time 1227.86 Seconds

(without alignments)
7596.186 Million cell updates/sec

Title: US-09-735-712-1

Sequence: 1 atgattcaagcaccgcaca.....attgtgacaatgtgttga 603

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba1:*

2: gb_ba2:*

3: gb_ba3:*

4: gb_in1:*

5: gb_in2:*

6: gb_in3:*

7: gb_om:*

8: gb_ov:*

9: gb_pat1:*

10: gb_pat2:*

11: gb_ph:*

12: gb_p11:*

13: gb_p12:*

14: gb_p13:*

15: gb_p14:*

16: em_ba1:*

17: em_ba2:*

18: em_fun:*

19: em_htgo_hum:*

20: em_htgo_inv:*

21: em_htgo_rod:*

22: em_htg_hum1:*

23: em_htg_hum2:*

24: em_htg_hum3:*

25: em_htg_hum4:*

26: em_htg_hum5:*

27: em_htg_hum6:*

28: em_htg_hum7:*

29: em_htg_hum8:*

30: em_htg_inv1:*

31: em_htg_inv2:*

32: em_htg_other:*

33: em_htg_rod:*

34: em_hum1:*

35: em_hum2:*

36: em_hum3:*

37: em_hum4:*

38: em_hum5:*

39: em_hum6:*

40: em_hum7:*

41: em_in:*

42: em_om:*

43: em_or:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	598.2	99.2	691	85	AB013103	AB013103 Homo sapi
2	598.2	99.2	694	89	AF237907	AF237907 Homo sapi
3	596.6	98.9	697	89	AF321127	AF321127 Homo sapi
4	157	26.0	138097	71	AC027787	AC027787 Homo sapi
5	157	26.0	147788	64	AC015840	AC015840 Homo sapi
6	157	26.0	161039	83	AP003127	AP003127 Homo sapi
7	157	26.0	166804	82	AP001034	AP001034 Homo sapi
8	153.4	25.4	138097	71	AC027787	AC027787 Homo sapi

C	9	153.4	25.4	161039	83	AP003127	Homo sapi
	10	115.6	19.2	166804	82	AP001034	Homo sapi
	11	88	14.6	1076	88	AF068288	Homo sapi
	12	86.4	14.3	916	85	AB013102	Homo sapi
	13	86.4	14.3	1619	89	AF237912	Homo sapi
	14	86.4	14.3	1669	9	AR035695	Sequence
	15	86.4	14.3	1669	9	AR083580	Sequence
	16	63.2	10.5	1646	97	HUMIERB	L35848 Homo sapien
	17	63.2	10.5	1661	9	A68627	Sequence 3
	18	63.2	10.5	1661	9	AR082014	Sequence 3
	19	63.2	10.5	1661	9	I25809	Sequence 3
	20	62.4	10.3	1222	7	SSR236932	Sus scrofa
	21	58.6	9.7	708	10	AX101306	Sequence
	22	58.6	9.7	1010	94	MUSECERB	J05019 Mouse mast
	23	55.6	9.2	899	94	AF237910	Mus muscu
	24	55.4	9.2	1163	94	AF237915	Mus muscu
	25	55	9.1	1474	10	I07272	Sequence 19
	26	55	9.1	1474	93	HSCD20	X12530 Human mRNA
	27	55	9.1	1476	10	I07277	Sequence 24
	28	55	9.1	1482	91	BC002807	Homo sapi
	29	55	9.1	1597	93	HSCD20A	X07203 Human mRNA
	30	54.8	9.1	518	94	AF280401	Mus muscu
	31	52.6	8.7	862	85	AB013104	Homo sapi
	32	52.6	8.7	930	89	AF286866	Homo sapi
	33	52.6	8.7	1131	89	AF212240	Homo sapi
	34	52.6	8.7	1152	89	AF237908	Homo sapi
	35	52.6	8.7	1289	89	AF142409	Homo sapi
	36	52.6	8.7	1488	9	AX017505	Sequence
	37	52.6	8.7	1679	9	AX017610	Sequence
	38	52	8.6	1182	89	AK000224	Homo sapi
	39	51.8	8.6	1146	97	HUMBL1YM	M27394 Human B-Lym
	40	51.4	8.5	1480	94	AF237914	Mus muscu
	41	50.8	8.4	1687	94	AF237909	Mus muscu
	42	49.4	8.2	969	89	AF201951	Homo sapi
	43	49.4	8.2	1012	85	AB026043	Homo sapi
	44	49.4	8.2	1228	89	AF237916	Homo sapi
	45	49.4	8.2	1228	94	AF237913	Mus muscu

ALIGNMENTS

RESULT 1
LOCUS AB013103 691 bp mRNA PRI 20-MAR-2001
DEFINITION Homo sapiens mRNA for MS4A5, complete cds.
ACCESSION AB013103
KEYWORDS MS4A5; CD20-like 2.

SOURCE Homo sapiens
ORGANISM Homo sapiens

REFERENCE 1 (sites)
AUTHORS Ishibashi, K., Suzuki, M., Sasaki, S. and Imai, M.
TITLE Identification of a new multigene four-transmembrane family (MS4A) related to CD20, HTM4 and beta subunit of the high-affinity Ige receptor

JOURNAL Gene 264 (1), 87-93 (2001)

REFERENCE 2 (sites)

AUTHORS Ishibashi, K., Sasaki, S. and Marumo, F.

TITLE Cloning of three CD20 homolog from human, putative calcium channels

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 691)

AUTHORS Ishibashi, K.

TITLE Direct Submission

JOURNAL Submitted (20-APR-1998) Kenichi Ishibashi, Tokyo Medical and Dental University, 2nd Internal Medicine; Yushima 1-5-45, Bunkyo, Tokyo

113-8519, Japan (E-mail: kishibashi.med2@med.tmd.ac.jp, Tel: 81-3-5803-5223, Fax: 81-3-5803-0132)

FEATURES

source location/Qualifiers

1..691

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/issue_type="testis"
52..654
/gene="MS4A5"
52..654
/gene="MS4A5"
/codon_start=1
/product="MS4A5"
/protein_id="BAB18739.1"
/db_xref="GI:11559214"
/translation="MDSSTHSEVFLVPPETIASVESTELSATPFSQSLQFLPA
RKMLIGTILQILFGIMTFSEGVLEFLTLKPYPRPFIISGVPFSGVLFINSGL
IAVRKTFEILILSRIMNPLSALGAIAGIILTFEFGIILIDONTICGSHONSCKAVT
VLEGLITLMPFSIIEFLISFSLISGLGHSECCDEQC"
BASE COUNT      195 a      142 c      117 g      237 t
ORIGIN

```

Query Match 99.2%; Score 598.2; DB 85; Length 691;
Best Local Similarity 99.5%; Pred. No. 1,4e-127;
Matches 600; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 atggaattcaagcaccgacacacagtcggtgttctglatlctccagaatactcgt 60
|||||
DB 52 ATGGATTCAAGCACCACGACACAGTCGCGTGTCTCTGATTTCTCCAGAAATCACTGCT 111
|||||
QY 61 tcggaatagatgctcagacagactttagcagcagccttcaactcaaaccccttgcaa 120
|||||
DB 112 TCAGAAATATGATGCACAGAACTTTCAGCAGCAGCTTTCACTCAAGCCCTTGCAA 171
|||||
QY 121 aaatatttcctagaanaaaagaaatccttagggactatccagatcctgttggaattatg 180
|||||
DB 172 AAATATTTCCTAGAAAATGAATCTTAGGAGCATTCAGATCCTGTTGGAAATATG 231
|||||
QY 181 acccttctcttgtagtattcctccttccacctgtgtaaacatataccaaggtttccc 240
|||||
DB 232 ACCTTTCTTTTGAGATTATCTTCTTCACTTTGTTAAACCATATCCAGGTTTCCC 291
|||||
QY 241 ttatatcttcttcagagatccatcttctgggcctcttctgcatatcttcgagacc 300
|||||
DB 292 TTTATATTTCTTCAGGATATCCATCTGGGGCTCTTTTGTCATTAATTCGGAGCC 351
|||||
QY 301 ttccctaatgcagtgtaaaagaaacacacagaaactctgataatattagccgataatg 360
|||||
DB 352 TTCTATATTCAGATGAGAAAGAAACACAGAAATCTGATATATGACCGGATATAG 411
|||||
QY 361 aatcttcttagtgcctcgtagagcaaatagctggaatcatcttctcacatttggttcatc 420
|||||
DB 412 AATTTCTTATGTCGCTGGAGCAATAGCTGAATCATCTTCCTCACATTTGGTTTCATC 471
|||||
QY 421 cttagtaaaactacatttctgtattctccacaaaatagtcagtgtaagcctgttact 480
|||||
DB 472 CTAGATCAAAACTACATTTGGTTATTTCTACCAAAATAGTCAAGTGAAGCGTTTACT 531
|||||
QY 481 gtccgtgcttggaatttgattatcatgatgacttcagcatcttattgaattatcatc 540
|||||
DB 532 GTCTGCTTCTTGGGAATTTGATTGATGATGATCTTCAGCANTATGAAATATATTCATT 591
|||||
QY 541 tctctgcttcttcaatttgggggtgcacactcagagagatgtgattgtgaacaatgttgt 600
|||||
DB 592 TCTCTGCTTCTTCAATTTGGGGTGCACCTAGAGAGATGTGATGTGTAACAATGTTGT 651
|||||
QY 601 tga 603
|||||
DB 652 TGA 654

```

RESULT 2
LOCUS AF237907 694 bp mRNA PRI 17-APR-2001
DEFINITION Homo sapiens MS4A5 protein mRNA, complete cds.
ACCESSION AF237907

VERSION AF237907.1 GI:13649400
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 694)
liang, Y. and Tedder, R.F.
TITLE Identification of a CD20-, Fc epsilon1alpha-beta-, and H19c4-related Gene
Family: Sixteen New MS4A Family Members Expressed in Human and
Mouse
JOURNAL Genomics 72 (2), 119-127 (2001)
REFERENCE 2 (bases 1 to 694)
AUTHORS liang, Y. and Tedder, R.F.
TITLE Direct Submission
JOURNAL Submitted (24-FEB-2000) Immunology, Duke Medical Center, Research
Dr., Durham, NC 27710, USA
FEATURES
source
1. 694
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS
55..657
/codon_start=1
/product="MS4A5 protein"
/protein_id="AAK37416.1"
/db_xref="GI:13649401"
/translation="MDSRAHSPVLPVPPETASRYESTELSATTFNSPDLKIFA
RKMKILGTIQLFGIMTFSEGVIFLFLKPPRPFFILSGIPMGSVLFINSGAFI
LAVKRTTETLILSRIMNLSALGAIILITFGLIDOWNICGYSHONSQCAVAT
VFLGLITLIMFSLIELFISLPSILGSHSDCDEOC"

BASE COUNT 198 a 143 c 116 g 237 t
ORIGIN

Query Match 99.2%; Score 598.2; DB 89; Length 694;
Best Local Similarity 99.5%; Pred. No. 1.4e-127;
Matches 600; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 atgattcaagcaccgacagctcgcggtttctgtgattcttcagagaatcagctgc 60
|||||
55 ATGATTCAAGCACCACGACAGCTCGGTGTTCTGTGATTTCTCCAGAAATCAGCTG 114
|||||
61 tcagaatagatccacagacttcagccagacacttcacactcaaaagcccttgcaa 120
|||||
115 TCAGAAATAGATCCACAGAACTTTCAGCCAGACCTTTCAACTCAAGCCCTTGCAA 174
|||||
121 aaattattgctagaanaaagaaatccttaggactataccagatcctgttggaaatag 180
|||||
175 AAATTATTGCTAGAAAATGAAAATCTTAGGACTATCCAGATCCTGTTTGAATTATG 234
|||||
181 acccttcttggaattatctccttcaactgtttaaaccatataccaaggtttccg 240
|||||
235 ACCTTTCTTTGGAGTATCTTCTTTCACCTTTGTTAAACCATATCCAAAGGTTTCC 294
|||||
241 ttataatcttcacagatatccatctcggggtcgtctgttggtaattccttgagcc 300
|||||
295 TTTATATTCTTCTCAGGATATCCTGCGGCTGCTGTTTGTCAATTAATCTGGAGCC 354
|||||
301 ttccataatgagtgagaagaagaaccacagaaactctgataatattgagcgaaataag 360
|||||
355 TTCTTAATTGCACTGAAAAGAAAACACAGAAACTCTGATTAATATGAGCCCAATTAATG 414
|||||
361 aactctcttagtgccctgagagaagaatgctggaatcattctctcaatttggttcacc 420
|||||
415 AATTTCCTTAGTCCCTGGAGCAATAGCTGGAATCAATCTCTCCACATTTGGTTTCAATC 474
|||||
421 cttagatcaaaactacatttgtgttlaattctacacaaatagtcagtgtaaggttact 480
|||||
475 CTGATGACAAACTACATTTGTGTTATCTTCACACAAATAGTCAAGTGAAGCTGTACT 534
|||||
481 gttcctgttcttgggaattttagatatacagacttcaagcaattatgaattatcatt 540
|||||
535 GTCCCTGTTCTTGGAATTTGATTACATGATGACATTTTCAGATTAATTAATTAATCATT 594
|||||

QY 541 tctctgcttcttcacatttgggggtgcccactcagagaggttgattgtgtgacaatgtgt 600
|||||
DB 595 TCCTGCTTTCATCAATTTGGGGTGCACACTCAGAGATGTGATTGTGAACAATGTTGT 654
|||||
QY 601 tga 603
|||||
DB 655 TGA 657

RESULT 3
AF321127
LOCUS AF321127 697 bp mRNA PRI 07-FEB-2001
DEFINITION Homo sapiens testis-expressed transmembrane-4 protein (TETM4) mRNA,
complete cds.
ACCESSION AF321127
VERSION AF321127.1 GI:12698681
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 697)
Hulet, M.D., Pagler, E., Hornby, J.R., Hogarth, P.M., Eyre, H.J.,
Baker, E., Crawford, J., Sutherland, G.R., Ohms, S.J. and Parish, C.R.
TITLE Isolation, tissue distribution, and chromosomal localization of a
novel testis-specific human four-transmembrane gene related to CD20
and Fc epsilon1alpha-beta
JOURNAL Biochem. Biophys. Res. Commun. 280 (1), 374-379 (2001)
MEDLINE 21092614
REFERENCE 2 (bases 1 to 697)
Hulet, M.D.
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (14-NOV-2000) Division of Immunology and Cell Biology,
John Curtin School of Medical Research, Mills Road, Canberra, ACT
2601, Australia
FEATURES
source
1. 697
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q12"
1. 697
/gene="TETM4"
57..659
/gene="TETM4"
/codon_start=1
/product="testis-expressed transmembrane-4 protein"
/protein_id="AAK01641.1"
/db_xref="GI:12698682"

BASE COUNT 199 a 142 c 119 g 237 t
ORIGIN

Query Match 98.9%; Score 596.6; DB 89; Length 697;
Best Local Similarity 99.3%; Pred. No. 3.2e-127;
Matches 599; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

1 atgattcaagcaccgacagctcgcggtttctgtgattcttcagagaatcagctgc 60
|||||
DB 57 ATGATTCAAGCACCACGACAGCTCGGTGTTCTGTGATTTCTCCAGAAATCAGCTGCT 116
|||||
QY 61 tcagaatagatccacagacttcagccagacacttcacactcaaaagcccttgcaa 120
|||||
DB 117 TCAGAAATAGATCCACAGAACTTTCAGCCAGACCTTTCAACTCAAGCCCTTGCAA 176
|||||
QY 121 aaattattgctagaanaaagaaatccttaggactataccagatcctgttggaaatag 180
|||||
DB 177 AAATTATTGCTAGAAAATGAAAATCTTAGGACTATCCAGATCCTGTTTGAATTATG 236
|||||

QY 181 acccttctcttgaggtatctctcttctacactgtttaaaccatataccaagttccc 240
 Db 237 ACCTTTCTCTTGGAGTATCTTCCTTTCACCTTGTAAACATATCCAGGTTCCC 296
 QY 241 ttataattcttcagagatataccatctcgtggctctgtttgttcatatccttgagcc 300
 Db 297 TTTATATTTCTTTCAGGATATCATCTCTGGGCTCTGTTTGTTCATTAATTCAGGCC 356
 QY 301 ttctcaattgcaatgaaaaaaccacagaaactcgtatataatgagccgaataatg 360
 Db 357 TTCTTAATTCAGATGAAAAAACCACAGAACTCTGATAATTTGAGCCCAATATATG 416
 QY 361 aatctctttagtgccctgagagacatagctgtaacatctctcctcaacttggttcact 420
 Db 417 AATTTCTTAGTGCCTCGGAGCATATGTCGATATCTCTCCTCACATTTGGTTTCATC 476
 QY 421 ctgagatcaaaactacattgtgtgttcttctacaaaataagtcagtgaagctgttact 480
 477 CTGAGATCAAAACTACATTTGCTGTTATTTCCACCAAAATAGTCAGTGTAGCGTTACT 536
 QY 481 gtctcttcttgaggaaatttgattacatgatacttcagcaatttgatattatcatt 540
 Db 537 GTCTCTGTGTTGGAAATTTGATTACATGATGATCTTCAGCATTTATGATTAATTCATT 596
 QY 541 tctctgaccttccaatttggggctgcacacacagagattgtatgttgacaaatgtgt 600
 Db 597 TCTCTGCTTTCTCAATTTGGGGTGCCACTCAGAGATTTGTGTAACATGTGTGT 656
 QY 601 tga 603
 Db 657 TGA 659

RESULT 4
 LOCUS AC027787/c
 DEFINITION Homo sapiens chromosome 15 clone RP11-196E16 map 15, LOW-PASS
 SEQUENCE SAMPLING.
 AC027787
 VERSION AC027787.2 GI:9845160
 KEYWORDS HTG: HTGS_PHASE0.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 138097)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Homo sapiens chromosome 15, clone RP11-196E16
 2 (bases 1 to 138097)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Baldwin,J., Barne,N., Bastien,V., Bede,F.,
 Boguski,K., Boukhgalter,B., Brown,A., Burkett,G.,
 Campione,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
 Collamore,A., Cooke,P., Deatellano,K., Dewar,K., Diaz,J.S.,
 Dodge,S., Domingo,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
 Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
 Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
 Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
 Klein,J., Laroque,K., Lamazares,R., Landers,T., Lehotzky,J.,
 Levine,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Margus,R.,
 McCarthy,M., McEwan,P., McGuirk,A., McKernan,K., McPheters,R.,
 Melidini,J., Menus,L., Mihova,T., Miranda,C., Mienna,V., Morrow,J.,
 Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
 O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
 Pisanil,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
 Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
 Testade,S., Theodore,J., Titrell,A., Travers,M., Triggillo,J.,
 Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
 Young,G., Zainou,J., Zimmer,A. and Zody,M.
 Direct Submission

JOURNAL COMMENT

Submitted (01-APR-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Aug 18, 2000 this sequence version replaced gi:7382629.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: MIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L9008
 Center clone name: 196_E_16

 * NOTE: This record contains 163 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.
 * 1
 * 664 763: gap of 100 bp
 * 764 1464: contig of 701 bp in length
 * 1465 1564: gap of 100 bp
 * 1565 2268: contig of 704 bp in length
 * 2269 2368: gap of 100 bp
 * 2369 3042: contig of 674 bp in length
 * 3043 3142: gap of 100 bp
 * 3143 3820: contig of 678 bp in length
 * 3821 3920: gap of 100 bp
 * 3921 4626: contig of 706 bp in length
 * 4627 4726: gap of 100 bp
 * 4727 5423: contig of 697 bp in length
 * 5424 5523: gap of 100 bp
 * 5524 6234: contig of 711 bp in length
 * 6235 6334: gap of 100 bp
 * 6335 7040: contig of 706 bp in length
 * 7041 7140: gap of 100 bp
 * 7141 7837: contig of 697 bp in length
 * 7838 7937: gap of 100 bp
 * 7938 8632: contig of 695 bp in length
 * 8633 8732: gap of 100 bp
 * 8733 9430: contig of 698 bp in length
 * 9431 9530: gap of 100 bp
 * 9531 10214: contig of 684 bp in length
 * 10215 10314: gap of 100 bp
 * 10315 10980: contig of 666 bp in length
 * 10981 11080: gap of 100 bp
 * 11081 11783: contig of 703 bp in length
 * 11784 11883: gap of 100 bp
 * 11884 12582: contig of 699 bp in length
 * 12583 12682: gap of 100 bp
 * 12683 13389: contig of 707 bp in length
 * 13390 13489: gap of 100 bp
 * 13490 14186: contig of 697 bp in length
 * 14187 14286: gap of 100 bp
 * 14287 14996: contig of 710 bp in length
 * 14997 15096: gap of 100 bp
 * 15097 15796: contig of 700 bp in length
 * 15797 15896: gap of 100 bp
 * 15897 16363: contig of 687 bp in length
 * 16364 16683: gap of 100 bp
 * 16684 17376: contig of 693 bp in length
 * 17377 17476: gap of 100 bp
 * 17477 18178: contig of 702 bp in length
 * 18179 18278: gap of 100 bp
 * 18279 18962: contig of 684 bp in length

```

* 18963 19062: gap of 100 bp
* 19063 19673: contig of 611 bp in length
* 19674 19773: gap of 100 bp
* 19774 20436: contig of 663 bp in length
* 20437 20536: gap of 100 bp
* 20537 21233: contig of 697 bp in length
* 21234 21333: gap of 100 bp
* 21334 22042: contig of 709 bp in length
* 22043 22142: gap of 100 bp
* 22143 22808: contig of 666 bp in length
* 22809 22908: gap of 100 bp
* 22909 23595: contig of 687 bp in length
* 23596 23695: gap of 100 bp
* 23696 24391: contig of 696 bp in length
* 24392 24491: gap of 100 bp
* 24492 25190: contig of 699 bp in length
* 25191 25290: gap of 100 bp
* 25291 25981: contig of 691 bp in length
* 25982 26081: gap of 100 bp
* 26082 26765: contig of 684 bp in length
* 26766 26865: gap of 100 bp
* 26866 27563: contig of 698 bp in length
* 27564 27663: gap of 100 bp
* 27664 28377: contig of 714 bp in length
* 28378 28477: gap of 100 bp
* 28478 29184: contig of 707 bp in length
* 29185 29284: gap of 100 bp
* 29285 29983: contig of 699 bp in length
* 29984 30083: gap of 100 bp
* 30084 30775: contig of 692 bp in length
* 30776 30875: gap of 100 bp
* 30876 31565: contig of 690 bp in length
* 31566 31665: gap of 100 bp
* 31666 32360: contig of 695 bp in length
* 32361 32460: gap of 100 bp
* 32461 33157: contig of 697 bp in length
* 33158 33257: gap of 100 bp
* 33258 33951: contig of 694 bp in length
* 33952 34051: gap of 100 bp
* 34052 34735: contig of 684 bp in length
* 34736 34835: gap of 100 bp
* 34836 35483: contig of 648 bp in length
* 35484 35583: gap of 100 bp
* 35584 36308: contig of 725 bp in length
* 36309 36408: gap of 100 bp
* 36409 37108: contig of 700 bp in length
* 37109 37208: gap of 100 bp
* 37209 37892: contig of 684 bp in length
* 37893 37992: gap of 100 bp
* 37993 38672: contig of 680 bp in length
* 38673 38772: gap of 100 bp
* 38773 39475: contig of 703 bp in length
* 39476 39575: gap of 100 bp
* 39576 40290: contig of 715 bp in length
* 40291 40390: gap of 100 bp
* 40391 41087: contig of 697 bp in length
* 41088 41187: gap of 100 bp
* 41188 41891: contig of 704 bp in length
* 41892 41991: gap of 100 bp
* 41992 42679: contig of 688 bp in length
* 42680 42779: gap of 100 bp
* 42780 43500: contig of 721 bp in length
* 43501 43600: gap of 100 bp
* 43601 44295: contig of 695 bp in length
* 44296 44395: gap of 100 bp
* 44396 45088: contig of 693 bp in length
* 45089 45188: gap of 100 bp
* 45189 45880: contig of 692 bp in length
* 45881 45980: gap of 100 bp
* 45981 46683: contig of 703 bp in length
* 46784 47481: gap of 100 bp
* 47482 47581: gap of 100 bp

```

```

Query Match          26.0%; Score 157; DB 71; Length 138097;
Best Local Similarity 91.7%; Pred. No. 4.8e-26;
Matches 166; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```

```

* 47582 48290: contig of 709 bp in length
* 48291 48390: gap of 100 bp
* 48391 49077: contig of 687 bp in length
* 49078 49177: gap of 100 bp
* 49178 49879: contig of 702 bp in length
* 49880 49979: gap of 100 bp
* 49980 50684: contig of 705 bp in length
* 50685 50784: gap of 100 bp
* 50785 51507: contig of 723 bp in length
* 51508 51607: gap of 100 bp
* 51608 52320: contig of 713 bp in length
* 52321 52420: gap of 100 bp
* 52421 53133: contig of 713 bp in length
* 53134 53233: gap of 100 bp
* 53234 53933: contig of 700 bp in length
* 53934 54033: gap of 100 bp
* 54034 54703: contig of 670 bp in length

```

```

QY 336 tctgataatattgagccgaataatgaatcttcttagtgccttagagagcaatagctggaat 335
|| |||||
Db 102667 TCAGATATATATGAGCCGATATGAAATTTCTTAGTCCCTGGAGCAATAGCTGGAAT 102608

QY 396 catcttcctcaatttggttcaatccttagatcaaaactaatattggttattctcaaca 455
|||||
Db 102607 CATTCCTCCACATTGGTTTCATCCTAGATCAAAACTAATGCTGCTATTCCTACCA 102548

QY 456 aatagtcagtcgaagcctgtactcgtcgttcttcttggaatttgattcatatgac 515
|||||
Db 102547 AATATGACATGTAAGGCTTTACTGTCTGCTCTTGTAAGATATGTCATTATATGAG 102488

QY 516 t 516
Db 102487 T 102487

```

```

RESULT 5
AC015840
LOCUS
DEFINITION
Homo sapiens clone RP11-24D1, WORKING DRAFT SEQUENCE, 5 unordered
pieces.
AC015840
AC015840.2 GI:7657730
VERSION
HTG: HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE
Homo sapiens
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 147788)
Birtten,B., Linton,L., Nusbaum,C. and Lander,E.
AUTHORS
Birtten,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE
Homo sapiens, clone RP11-24D1
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 147788)
Birtten,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barina,N., Beckert,R., Boguslavsky,L., Boukhalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearellano,K., Dewar,K., Domingo,M., Donegan,L., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howard,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,D.,
Lehoczy,J.C., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McQuirk,A., McKernan,K., McLaughlin,D., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Stojanovic,N., Subramanian,A., Talmas,J.,
Stefaye,S., Tittel,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

```

```

TITLE
JOURNAL
Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

```



```

137320 143276 contig of 5957 bp in length
143377 148922 contig of 5567 bp in length
149023 154159 contig of 5137 bp in length
154260 157686 contig of 3427 bp in length
157787 159234 contig of 1448 bp in length
159335 161039 contig of 1705 bp in length.
*
* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 42322: contig of 42322 bp in length
*
42323 42422: gap of 100 bp
*
42423 63418: contig of 20996 bp in length
*
63419 63518: gap of 100 bp
*
63519 77969: contig of 14451 bp in length
*
77970 78069: gap of 100 bp
*
78070 93182: contig of 15113 bp in length
*
93183 93282: gap of 100 bp
*
93283 104093: contig of 10811 bp in length
*
104094 104193: gap of 100 bp
*
104194 115702: contig of 11509 bp in length
*
115703 115802: gap of 100 bp
*
115803 125183: contig of 9381 bp in length
*
125184 125283: gap of 100 bp
*
125284 132558: contig of 7275 bp in length
*
132559 132658: gap of 100 bp
*
132659 137219: contig of 4561 bp in length
*
137220 137319: gap of 100 bp
*
137320 143276: contig of 5957 bp in length
*
143277 143376: gap of 100 bp
*
143377 148922: contig of 5546 bp in length
*
148923 149022: gap of 100 bp
*
149023 154159: contig of 5137 bp in length
*
154160 154259: gap of 100 bp
*
154260 157686: contig of 3427 bp in length
*
157687 157786: gap of 100 bp
*
157787 159234: contig of 1448 bp in length
*
159235 159334: gap of 100 bp
*
159335 161039: contig of 1705 bp in length.
*
Location/Qualifiers
source
1. 161039
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-710G6"
1. 42322
/misc_feature
/note="assembly_fragment"
42423 63418
/note="assembly_fragment"
63519 77969
/note="assembly_fragment"
78070 93182
/note="assembly_fragment"
93283 104093
/note="assembly_fragment"
104194 115702
/note="assembly_fragment"
115803 125183
/note="assembly_fragment"
125284 132558
/note="assembly_fragment"
132659 137219
/note="assembly_fragment clone_end:SP6 vector_side:left"
137320 143276
/note="assembly_fragment"
143377 148922
/note="assembly_fragment clone_end:T7 vector_side:left"
149023 154159

```

	misc_feature	/note="assembly_fragment" 154260 .157686	
	misc_feature	/note="assembly_fragment" 157787 .159234	
	misc_feature	/note="assembly_fragment" 159335 .161039	
		/note="assembly_fragment"	
BASE COUNT	49155 a 31166 c 31523 g 47795 t	1400 others	
ORIGIN			
Query Match	26.0%; Score 157; DB 83;	Length 161039;	
Best Local Similarity	91.7%;	Pred. NO. 4.9e-26;	
Matches 166; Conservative	0; Mismatches 15;	Indels 0; Gaps 0;	
OY	336	tctgataatatgaagccagaataatgaaccttccttagtgcgccgtgagagcaaatagtctgaaat	395
Db	45928	TTCAGTAAATATTGAGCCGAATATATAAATTTCTTAGTGCCCGGGACCAATAGCTGSAAAT	45987
OY	396	cattcctcatcaatttgttlccaacctgatcalcaaaaactacaatttgttgtaattccacca	455
Db	45988	CATTCCTCCTCACATTTGGTTGCATCTGTGANCACAANAAGCATTTGTGGTTATTTCCACACA	46047
OY	456	aaatagtcagtgaaggctgttaactgtccccgtctcttggtggaatttgattacattgaatgac	515
Db	46048	AAAATGTCAGGTAAAGCGCTGTACTGTCTCCTGTTCTTGGTAAGTATGTCATTPTTAGAG	46107
OY	516	t 516	
Db	46108	T 46108	

RESULT	7
AP001034/c	
LOCUS	AP001034 166804 bp DNA HTG 14-NOV-2000
DEFINITION	Homo sapiens chromosome 11 clone RP11-729B4 map 11q12, WORKING
ACCESSION	DRAFT SEQUENCE, 19 unordered pieces.
AP001034	
VERSION	AP001034.4 GI:11176992
KEYWORDS	HTG, HTGS, PHASE1, HTGS, DRAFT.
SOURCE	Homo sapiens DNA, clone:RP11-729B4.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota: Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 166804)
TITLE	Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
JOURNAL	Homo sapiens 166,804 genomic DNA of 11q12
REFERENCE	Published Only in Database (2000) In press
AUTHORS	2 (bases 1 to 166804)
TITLE	Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
JOURNAL	Direct Submission
COMMENT	Submitted (06-JAN-2000) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Sehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT	On Nov 14, 2000 this sequence version replaced gi:8117704.
COMMENT	----- Genome Center -----
COMMENT	Center: RIKEN Genomic Sciences Center(GSC)
COMMENT	Center code: RIKEN
COMMENT	Web site: http://hgp.gsc.riken.go.jp/
COMMENT	Contact: hattori@gsc.riken.go.jp
COMMENT	----- Project Information -----
COMMENT	Center project name: HumDrat11
COMMENT	Center clone name: RP11-729B4
COMMENT	----- Summary Statistics -----
COMMENT	Sequencing vector: PCR products, 100% of reads
COMMENT	Chemistry: Dye-terminator ET-primersham, 100% of reads
COMMENT	Assembly program: Phrap, version 0.990329
COMMENT	Consensus quality: 162254 bases at least Q40
COMMENT	Consensus quality: 163854 bases at least Q30

Consensus quality: 164544 bases at least Q20
 Insert size: 165004; sum-of-contigs
 Quality coverage: 9.28x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 19 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs 'N', but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1 22398 contig of 22398 bp in length
 22499 45238 contig of 22740 bp in length
 45339 64278 contig of 18940 bp in length
 64379 77579 contig of 13201 bp in length
 77680 93718 contig of 16039 bp in length
 93819 108057 contig of 14239 bp in length
 108158 117006 contig of 8849 bp in length
 117107 126302 contig of 9196 bp in length
 126403 133479 contig of 7077 bp in length
 133580 140425 contig of 6846 bp in length
 140526 144653 contig of 4128 bp in length
 144754 149354 contig of 4601 bp in length
 149455 154522 contig of 5068 bp in length
 154623 155919 contig of 1297 bp in length
 156020 159504 contig of 3485 bp in length
 159605 159679 contig of 75 bp in length
 159780 162418 contig of 2639 bp in length
 162519 165383 contig of 2865 bp in length
 165484 166804 contig of 1321 bp in length.

NOTE: This is a 'working draft' sequence. It currently consists of 19 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of 'N', but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 22398 contig of 22398 bp in length
 22399 22498: gap of 100 bp
 22499 45238: contig of 22740 bp in length
 45339 45338: gap of 100 bp
 45339 64278: contig of 18940 bp in length
 64279 64378: gap of 100 bp
 64379 77579: contig of 13201 bp in length
 77580 77679: gap of 100 bp
 77680 93718: contig of 16039 bp in length
 93719 93818: gap of 100 bp
 93819 108057: contig of 14239 bp in length
 108058 108157: gap of 100 bp
 108158 117006: contig of 8849 bp in length
 117007 117106: gap of 100 bp
 117107 126302: contig of 9196 bp in length
 126303 126402: gap of 100 bp
 126403 133479: contig of 7077 bp in length
 133480 133579: gap of 100 bp
 133580 140425: contig of 6846 bp in length
 140426 140525: gap of 100 bp
 140526 144653: contig of 4128 bp in length
 144654 144753: gap of 100 bp
 144754 149354: contig of 4601 bp in length
 149355 149454: gap of 100 bp
 149455 154522: contig of 5068 bp in length
 154523 154622: gap of 100 bp
 154623 155919: contig of 1297 bp in length
 155920 156019: gap of 100 bp
 156020 159504: contig of 3485 bp in length
 159505 159604: gap of 100 bp
 159605 159679: contig of 75 bp in length
 159680 159779: gap of 100 bp
 159780 162418: contig of 2639 bp in length
 162419 162518: gap of 100 bp
 162519 165383: contig of 2865 bp in length

* 165384 165483: gap of 100 bp
 * 165484 166804: contig of 1321 bp in length.
 FEATURES
 source
 1.166804
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q12"
 /clone="RP11-729B4"

misc.feature
 1.22398
 /note="assembly-fragment"
 misc.feature
 22499..45238
 /note="assembly-fragment"
 misc.feature
 45339..64278
 /note="assembly-fragment"
 misc.feature
 64379..77579
 /note="assembly-fragment"
 misc.feature
 77680..93718
 /note="assembly-fragment"
 misc.feature
 93819..108057
 /note="assembly-fragment"
 misc.feature
 108158..117006
 /note="assembly-fragment"
 misc.feature
 117107..126302
 /note="assembly-fragment"
 misc.feature
 126403..133479
 /note="assembly-fragment"
 misc.feature
 133580..140425
 /note="assembly-fragment"
 misc.feature
 140526..144653
 /note="assembly-fragment"
 misc.feature
 144754..149354
 /note="assembly-fragment"
 misc.feature
 149455..154522
 /note="assembly-fragment"
 misc.feature
 154623..155919
 /note="assembly-fragment"
 misc.feature
 156020..159504
 /note="assembly-fragment"
 misc.feature
 159605..159679
 /note="assembly-fragment clone_end:17 vector_side:left"
 misc.feature
 159780..162418
 /note="assembly-fragment"
 misc.feature
 162519..165383
 /note="assembly-fragment"
 misc.feature
 165484..166804
 /note="assembly-fragment"
 BASE COUNT 50706 a 31523 c 32236 g 50478 t 1801 others
 ORIGIN

Query Match 26.0%; Score 157; DB 82; Length 166804;
 Best Local Similarity 91.7%; Pred. No. 4.9e-26;
 Matches 166; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 336 tctgataatattgagccgaataatgaattctcttagtccttgagagcaatgctgnaat 395
 || |||||
 DB 7669 TCAGATAATATGAGCGAATATGAAATTTCTTAGGCCCTGGAGCAATACCTGGAAT 7610
 QY 396 catctccctcaatttggttcatctctagatcaaaactacatttggttatcttaccac 455
 |||||
 DB 7609 CATTTCCTCCTACATTTGGTTTCATCTAGATCAAACTACATTTGGTTATCTCACC 7550
 QY 456 aatatgcaatgtaagagcgttctctcgtctctgggaatttatcatatgtagac 515
 |||||
 DB 7549 AATATGACAGTGAAGCGTGTACTGCTGTTCTTGATGATGTGCATTATAGAG 7490
 QY 516 t 516
 |
 DB 7489 T 7489

RESULT 8

AC027787
LOCUS AC027787 138097 bp DNA HTG 18-AUG-2000
DEFINITION Homo sapiens chromosome 15 clone RP11-196E16 map 15, LOW-PASS
SEQUENCE SAMPLING.
AC027787
VERSION AC027787.2 GI:9845160
KEYWORDS HTG; HTGS_PHASE0.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Birren, B., Linton, L., Nusbaum, C., and Lander, E.
TITLE 1 (bases 1 to 138097)
JOURNAL Homo sapiens chromosome 15, clone RP11-196E16
REFERENCE
AUTHORS unpublished
TITLE 2 (bases 1 to 138097)
JOURNAL Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G.,
Campolino, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., Deatellano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., Laroque, K., Lamazares, R., Landers, T., Lehoczy, J.,
Levine, R., Lien, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
McCarthy, M., McKean, P., McGuirk, A., McKernan, K., McPheeters, R.,
Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mlenka, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange, Thoman, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Testage, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (01-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 18, 2000 this sequence version replaced gi:7382629.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L9908
Center clone name: 196_E_16

* NOTE: This record contains 163 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1
* 664 763: gap of 663 bp in length
* 764 1464: contig of 701 bp in length
* 1465 1564: gap of 100 bp
* 1565 2268: contig of 704 bp in length
* 2269 2368: gap of 100 bp
* 2369 3042: contig of 674 bp in length
* 3043 3142: gap of 100 bp
* 3143 3820: contig of 678 bp in length

* 3821 3920: gap of 100 bp
* 3921 4626: contig of 706 bp in length
* 4627 4726: gap of 100 bp
* 4727 5423: contig of 697 bp in length
* 5424 5523: gap of 100 bp
* 5524 6234: contig of 711 bp in length
* 6235 6334: gap of 100 bp
* 6335 7040: contig of 706 bp in length
* 7041 7140: gap of 100 bp
* 7141 7837: contig of 697 bp in length
* 7838 7937: gap of 100 bp
* 7938 8632: contig of 695 bp in length
* 8633 8732: gap of 100 bp
* 8733 9430: contig of 698 bp in length
* 9431 9530: gap of 100 bp
* 9531 10214: contig of 684 bp in length
* 10215 10314: gap of 100 bp
* 10315 10980: contig of 666 bp in length
* 10981 11080: gap of 100 bp
* 11081 11783: contig of 703 bp in length
* 11784 11883: gap of 100 bp
* 11884 12582: contig of 699 bp in length
* 12583 13389: contig of 707 bp in length
* 13390 13489: gap of 100 bp
* 13490 14186: contig of 697 bp in length
* 14187 14286: gap of 100 bp
* 14287 14996: contig of 710 bp in length
* 14997 15096: gap of 100 bp
* 15097 15796: contig of 700 bp in length
* 15797 15896: gap of 100 bp
* 15897 16583: contig of 687 bp in length
* 16584 16683: gap of 100 bp
* 16684 17376: contig of 693 bp in length
* 17377 17476: gap of 100 bp
* 17477 18178: contig of 702 bp in length
* 18179 18278: gap of 100 bp
* 18279 18962: contig of 684 bp in length
* 18963 19062: gap of 100 bp
* 19063 19673: contig of 611 bp in length
* 19674 19773: gap of 100 bp
* 19774 20436: contig of 663 bp in length
* 20437 20536: gap of 100 bp
* 20537 21233: contig of 697 bp in length
* 21234 21333: gap of 100 bp
* 21334 22042: contig of 709 bp in length
* 22043 22142: gap of 100 bp
* 22143 22808: contig of 666 bp in length
* 22809 22908: gap of 100 bp
* 22909 23595: contig of 687 bp in length
* 23596 23695: gap of 100 bp
* 23696 24391: contig of 696 bp in length
* 24392 24491: gap of 100 bp
* 24492 25190: contig of 699 bp in length
* 25191 25290: gap of 100 bp
* 25291 25981: contig of 691 bp in length
* 25982 26081: gap of 100 bp
* 26082 26765: contig of 684 bp in length
* 26766 26865: gap of 100 bp
* 26866 27563: contig of 698 bp in length
* 27564 27663: gap of 100 bp
* 27664 28377: contig of 714 bp in length
* 28378 28477: gap of 100 bp
* 28478 29184: contig of 707 bp in length
* 29185 29284: gap of 100 bp
* 29285 29983: contig of 699 bp in length
* 29984 30083: gap of 100 bp
* 30084 30775: contig of 692 bp in length
* 30776 30875: gap of 100 bp
* 30876 31565: contig of 690 bp in length
* 31566 31665: gap of 100 bp
* 31666 32360: contig of 695 bp in length
* 32361 32460: gap of 100 bp

```

* 32461 33157: contig of 697 bp in length
* 33158 33257: gap of 100 bp
* 33258 33951: contig of 694 bp in length
* 33952 34051: gap of 100 bp
* 34052 34735: contig of 684 bp in length
* 34736 34835: gap of 100 bp
* 34836 35483: contig of 648 bp in length
* 35484 35583: gap of 100 bp
* 35584 36308: contig of 725 bp in length
* 36309 36408: gap of 100 bp
* 36409 37108: contig of 700 bp in length
* 37109 37208: gap of 100 bp
* 37209 37892: contig of 684 bp in length
* 37893 37992: gap of 100 bp
* 37993 38672: contig of 680 bp in length
* 38673 38772: gap of 100 bp
* 38773 39475: contig of 703 bp in length
* 39476 39575: gap of 100 bp
* 39576 40290: contig of 715 bp in length
* 40291 40390: gap of 100 bp
* 40391 41087: contig of 697 bp in length
* 41088 41187: gap of 100 bp
* 41188 41891: contig of 704 bp in length
* 41892 41991: gap of 100 bp
* 41992 42679: contig of 688 bp in length
* 42680 42779: gap of 100 bp
* 42780 43500: contig of 721 bp in length
* 43501 43600: gap of 100 bp
* 43601 44295: contig of 695 bp in length
* 44296 44395: gap of 100 bp
* 44396 45088: contig of 693 bp in length
* 45089 45188: gap of 100 bp
* 45189 45880: contig of 692 bp in length
* 45881 45980: gap of 100 bp
* 45981 46683: contig of 703 bp in length
* 46684 46783: gap of 100 bp
* 46784 47481: contig of 698 bp in length
* 47482 47581: gap of 100 bp
* 47582 48290: contig of 709 bp in length
* 48291 48390: gap of 100 bp
* 48391 49077: contig of 687 bp in length
* 49078 49177: gap of 100 bp
* 49178 49879: contig of 702 bp in length
* 49880 49979: gap of 100 bp
* 49980 50684: contig of 705 bp in length
* 50685 50784: gap of 100 bp
* 50785 51507: contig of 723 bp in length
* 51508 51607: gap of 100 bp
* 51608 52320: contig of 713 bp in length
* 52321 52420: gap of 100 bp
* 52421 53133: contig of 713 bp in length
* 53134 53233: gap of 100 bp
* 53234 53933: contig of 700 bp in length
* 53934 54033: gap of 100 bp
* 54034 54703: contig of 670 bp in length

```

Query Match 25.4%; Score 153.4; DB 71; Length 138097;
 Best Local Similarity 96.3%; Pred. No. 3.2e-25;
 Matches 157; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

QY 1 atgattcaagcaccgacacagtcggtcttctgattctcctcagaataactgct 60
    |||||||
DB 109596 ATGGATTCAAGCACCACGACAGTCGGTCTTCTGATTTCCTCCAGAAATCAGT 109655
    |||||||

QY 61 tcgaataatagctccacagacttcgacacacacttcaactcaaaccccttgcaa 120
    |||||||
DB 109656 TCGAATATAGTCCACAGACTTCAGCCACGACCTTTCACTCAATCAAGCCCTTGCAA 109715
    |||||||

QY 121 aaattatttctagaanaaatccttagagcattccaga 163
    |||||||
DB 109716 AAATTTATTTCTAGAAATAATGAAATCTTAGGGGTAGTAGTAAGA 109758
    |||||||

```

```

RESULT 9
AP003127/c
LOCUS
DEFINITION
Homosapiens chromosome 11 clone RP11-71066 map 11q, WORKING DRAFT
SEQUENCE, 15 unordered pieces.
AP003127
ACCESSION
AP003127.1 GI:12597183
VERSION
HTG: HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 161039)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens 161,039 genomic DNA of 11q
Published only in Database (2001) In press
2 (bases 1 to 161039)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (26-JAN-2001) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suenho-cho, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: HumDrat11
Center clone name: RP11-71066
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 157456 bases at least Q40
Consensus quality: 158960 bases at least Q30
Consensus quality: 159457 bases at least Q20
Insert size: 159639; sum-of-contigs
Quality coverage: 8.61x in Q20 bases; sum-of-contigs

```

NOTE: This is a 'working draft' sequence. It currently consists of 15 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs 'N', but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

```

1 42322 contig of 42322 bp in length
42423 63418 contig of 20996 bp in length
63519 77969 contig of 14451 bp in length
78070 93182 contig of 15113 bp in length
93283 104093 contig of 10811 bp in length
104194 115702 contig of 11509 bp in length
115803 125183 contig of 9581 bp in length
125284 132558 contig of 7275 bp in length
132659 137219 contig of 4561 bp in length
137320 143276 contig of 5957 bp in length
143377 148922 contig of 5546 bp in length
149023 154159 contig of 5137 bp in length
154260 157686 contig of 3427 bp in length
157787 159234 contig of 1448 bp in length
159335 161039 contig of 1705 bp in length.

```

NOTE: This is a 'working draft' sequence. It currently consists of 15 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs 'N', but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will

```

* be preserved.
* 1 42322: contig of 42322 bp in length
* 42323 42422: gap of 100 bp
* 42423 63418: contig of 20996 bp in length
* 63419 63518: gap of 100 bp
* 63519 77969: contig of 14451 bp in length
* 77970 78069: gap of 100 bp
* 78070 93182: contig of 15113 bp in length
* 93183 93282: gap of 100 bp
* 93283 104093: contig of 10811 bp in length
* 104094 104193: gap of 100 bp
* 104194 115702: contig of 11509 bp in length
* 115703 115802: gap of 100 bp
* 115803 125183: contig of 9381 bp in length
* 125184 125283: gap of 100 bp
* 125284 132358: contig of 7275 bp in length
* 132359 132658: gap of 100 bp
* 132659 137219: contig of 4561 bp in length
* 137220 137319: gap of 100 bp
* 137320 143276: contig of 5957 bp in length
* 143277 143376: gap of 100 bp
* 143377 148922: contig of 5546 bp in length
* 148923 149022: gap of 100 bp
* 149023 154159: contig of 5137 bp in length
* 154160 154259: gap of 100 bp
* 154260 157686: contig of 3427 bp in length
* 157687 157786: gap of 100 bp
* 157787 159234: contig of 1448 bp in length
* 159235 159334: gap of 100 bp
* 159335 161039: contig of 1705 bp in length.
* Location/Qualifiers
* 1.161039
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /chromosome="11"
  /map="11q"
  /clone="RP11-710G6"
  1.42322
    misc_feature
      /note="assembly_fragment"
      42423..63418
    misc_feature
      /note="assembly_fragment"
      63519..77969
    misc_feature
      /note="assembly_fragment"
      78070..93182
    misc_feature
      /note="assembly_fragment"
      93283..104093
    misc_feature
      /note="assembly_fragment"
      104194..115702
    misc_feature
      /note="assembly_fragment"
      115803..125183
    misc_feature
      /note="assembly_fragment"
      125284..132358
    misc_feature
      /note="assembly_fragment"
      132659..137219
    misc_feature
      /note="assembly_fragment clone_end:SP6 vector_side:left"
      137320..143276
    misc_feature
      /note="assembly_fragment"
      143377..148922
    misc_feature
      /note="assembly_fragment clone_end:T7 vector_side:left"
      149023..154159
    misc_feature
      /note="assembly_fragment"
      154260..157686
    misc_feature
      /note="assembly_fragment"
      157787..159234
    misc_feature
      /note="assembly_fragment"
      159335..161039
    misc_feature
      /note="assembly_fragment"
      161039..161039
  misc_feature
    /note="assembly_fragment"
    49155 a 31166 c 31523 g 47795 t 1400 others
  ORIGIN

```

```

Query Match 25.4%; Score 153.4; DB 83; Length 161039;
Best Local Similarity 96.3%; Pred. No. 3.3e-25;

```

```

Matches 157; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 atgattcaagcaccgacacagtcggtgtttctgttattcttcagaatactcgt 60
|||||
Db 149618 ATGATTCAAGCACCGACACAGTCGGTGTTCGTGATTTCCTCCAGAAATCACTGCT 149559
|||||
QY 61 tcaaatatgacccacagacttcagcagcacttcacacttcacaaagcccttcga 120
|||||
Db 149558 TCAGATATATGATCCACAGACTTTCAGCCAGACCTTTCAACTCAAGCCCTTCGAA 149499
|||||
QY 121 aaattattgctagaanaatcgaatcttagggactatccaga 163
|||||
Db 149498 AAATTTATTTCTAGAAAATGAATGAAATCTTAGGGGTAGTAGA 149456
|||||

RESULT 10
AP001034
LOCUS
DEFINITION
Homo sapiens chromosome 11 clone RP11-729B4 map 11q12, WORKING
DRAFT SEQUENCE. 19 unordered pieces.
ACCESSION
AP001034 GI:11176992
VERSION
AP001034.4
KEYWORDS
HTG; HTGS; PHASE1; HTGS; DRAFT.
SOURCE
Homo sapiens DNA, clone:RP11-729B4.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 166804)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens 166,804 genomic DNA of 11q12.
Published Only in Database (2000) In press
2 (bases 1 to 166804)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (06-JAN-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@psc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
On Nov 14, 2000 this sequence version replaced gi:8117704.
COMMENT
----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@psc.riken.go.jp
----- Project Information
Center Project name: HumDrafl11
Center clone name: RP11-729B4
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 162254 bases at least Q40
Consensus quality: 163858 bases at least Q30
Consensus quality: 164544 bases at least Q20
Insert size: 165004; sum-of-ctrls
Insert coverage: 9.28x in Q20 bases; sum-of-ctrls
-----
NOTE: This is a 'working draft' sequence. It currently consists of
19 contigs. The true order of the pieces is not known and the
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved
1 22398 contig of 22398 bp in length
22499 45238 contig of 22740 bp in length
45339 64278 contig of 18940 bp in length
64379 77579 contig of 13201 bp in length
77680 93718 contig of 16039 bp in length
93819 108057 contig of 14239 bp in length

```



```

source
1. .1076
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="dendritic cell"
196. .789
/notes="possibly a receptor or membrane associated protein"
/codon_start=1
/product="HDCME31P"
/protein_id="AF65507.1"
/db_xref="GI:7643784"
/translation="MAVISHLMKIQEFLKGPVLYGVQVLLTALMSLMTMC
MASNTGSPNLSVHIGYITWGYFLISGLSIAAGIRRTGLVRSGLMSVLA
ASGILINTSLAFYSPHPCNYGNSNCHGMSILMDGAVLLLSVLEFCIAVSL
SAFGCVLCTPGVLLILPSHMAETASPTPLNEY"

BASE COUNT      318 a      210 c      231 g      317 t
ORIGIN

Query Match      14.6%; Score 88; DB 88; Length 1076;
Best Local Similarity 51.3%; Pred. No. 2.1e-10;
Matches 233; Conservative 0; Mismatches 215; Indels 6; Gaps 1;

119 aaaaattatttgcagaataaataaatacttaaggactatccagatcccttggaaat 178
|||||
236 AGAAGTCTTGAAAGGGAACCCAAAGCTTGCGGTTGTGACAGATTCTGACCTCGA 295
|||||

179 tgaccttcttggaggtatcttcttccattcacttctgttaaacatccaaagttc 238
|||||
296 TGAGCCTTACAGTGGGAATACATGATGTGTATGCGATCTTAATCTTATGAAATACC 355
|||||

239 ccttataattcttcagatatccattctgggctctgttcttgcatttaattcggag 298
|||||
356 CTATTTCGGTCATATGCGGTACACATTTGGGGGTGACGATGTTATTTATTTCAGAT 415
|||||

299 ccttcccaattgcagtgaaagaaacacacagaactctgataatactgaagcgaata 388
|||||
416 CTTTGTCAATTGACGAGGAAATTAAGAACTCAAAAGGCTGCGGATGATCTAGGAA 475
|||||

359 tgaactcttcttaagtcctcctgaagcaatagctgaatcattccctcacaattggttca 418
|||||
476 TGAATATCACACACTGTGATCGCTGCATCAGGATCTTAATACACACTTTAGCTTGG 535
|||||

419 tccatagat-----caaaactacatttgytattctcaccacaatagtcagtgtaag 472
|||||
536 CGTTTATTCATTCACACCTTACTGTACTGATGCACTATGCACTCAATAATTTGTCATG 595
|||||

473 ctgttactgcctgttcttgggaatttgattacattgacttgcagcattatggaat 532
|||||
596 GGACTATGTCATCTTAATAGGCTGTGATGGCATGGTCTCTTAAGTGTCTGGAAAT 655
|||||

533 tattcattctcgccttctcatttggggtg 566
|||||
656 TCTGCATTCGTCTGTCCCTCTCTGCTTTGGATG 689
|||||

RESULT 12
AB013102      916 bp      mRNA      PRI      20-MAR-2001
LOCUS      Homo sapiens mRNA for MS4A4, complete cds.
DEFINITION
AB013102.1 GI:11559211
VERSION
MS4A4; CD20-like 1.
KEYWORDS
SOURCE      Homo sapiens fetus cDNA to mRNA.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
REFERENCE
AUTHORS      Ishibashi,K., Suzuki,M., Sasaki,S. and Imai,M.
TITLE      Identification of a new multigene four-transmembrane family (MS4A)
related to CD20, HTM4 and beta subunit of the high-affinity Ige
receptor
JOURNAL      Gene 264 (1), 87-93 (2001)
MEDLINE      2114397

```

```

REFERENCE      2 (bases 1 to 916)
AUTHORS      Ishibashi,K.
TITLE      Direct Submission
JOURNAL      Submitted (20-APR-1998) Kenichi Ishibashi, Tokyo Medical and Dental
University, 2nd Internal Medicine; Yushima 1-5-45, Bunkyo, Tokyo
113-8519, Japan (E-mail:kishibashi.med2@med.tmd.ac.jp,
Tel:81-3-5803-5223, Fax:81-3-5803-0132)
location/Qualifiers
FEATURES
source
1. 916
/organism="Homo sapiens"
/db_xref="taxon:9606"
/dev_stage="fetus"
79. .696
/gene="MS4A4"
79. .696
/gene="MS4A4"
/gene="MS4A4"
/product="MS4A4"
/codon_start=1
/protein_id="BAB18738.1"
/db_xref="GI:11559212"
/translation="MDVPGNMAVTHSHLMKIQEFLKGPVLYGVQVLLTALMSL
SMGRTMCAASNTGSPNLSVHIGYITWGYFLISGLSIAAGIRRTGLVRSGLM
NTTSSVLAASGILINTSLAFYSPHPCNYGNSNCHGMSILMDGAVLLLSVL
EFCIAVSLSAFCKVLCCTPGVLLILPSHMAETASPTPLNEY"

BASE COUNT      255 a      193 c      199 g      269 t
ORIGIN

Query Match      14.3%; Score 86.4; DB 85; Length 916;
Best Local Similarity 51.1%; Pred. No. 4.9e-10;
Matches 232; Conservative 0; Mismatches 216; Indels 6; Gaps 1;

119 aaaaattatttgcagaataaataaatacttaaggactatccagatcccttggaaat 178
|||||
143 AGAAGTCTTGAAAGGGAACCCAAAGCTTGCGGTTGTGACAGATTCTGACCTCGA 202
|||||

179 tgaccttcttggaggtatcttcttccattcacttctgttaaacatccaaagttc 238
|||||
203 TGAGCCTTACAGTGGGAATACATGATGTGTATGCGATCTTAATCTTATGAAATACC 262
|||||

239 ccttataattcttcagatatccattctgggctctgttcttgcatttaattcggag 298
|||||
263 CTATTTCGGTCATATGCGGTACACATTTGGGGGTGACGATGTTATTTATTTCAGAT 322
|||||

299 ccttcccaattgcagtgaaagaaacacacagaactctgataatactgaagcgaata 358
|||||
323 CTTTGTCAATTGACGAGGAAATTAAGAACTCAAAAGGCTGCGGATGATCTAGGAA 382
|||||

359 tgaactcttcttaagtcctcctgaagcaatagctgaatcattccctcacaattggttca 418
|||||
383 TGAATATCACACACTGTGATCGCTGCATCAGGATCTTAATACACACTTTAGCTTGG 442
|||||

419 tccatagat-----caaaactacatttgytattctcaccacaatagtcagtgtaag 472
|||||
443 CGTTTATTCATTCACACCTTACTGTACTGATGCACTATGCACTCAATAATTTGTCATG 502
|||||

473 ctgttactgcctgttcttgggaatttgattacattgacttgcagcattatggaat 532
|||||
503 GGACTATGTCATCTTAATAGGCTGTGATGGCATGGTCTCTTAAGTGTCTGGAAAT 562
|||||

533 tattcattctcgccttctcatttggggtg 566
|||||
563 TCTGCATTCGTCTGTCCCTCTCTGCTTTGGATG 596
|||||

RESULT 13
AF237912      1619 bp      mRNA      PRI      17-APR-2001
LOCUS      Homo sapiens MS4A4 protein mRNA, complete cds.
DEFINITION
AF237912
ACCESSION      AF237912.1 GI:13655456
VERSION
KEYWORDS
SOURCE      human.

```


ORIGIN

Query Match 14.3%; Score 86.4; DB 9; Length 1669;
Best Local Similarity 51.1%; Pred. No. 5.1e-10;
Matches 232; Conservative 0; Mismatches 216; Indels 6; Gaps 1;

```
QY 119 aaaaattattgtcagaanaaaatgaagaactctaggaactatccagatccgttttgaatta 178
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 325 AGAAGTCTTGAGAGGAGAGAACCCAAAGTCTTGSGGTGTGACAGATTCTGACTGCCCTGA 384
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 179 tgaactttctcttgaggtatcttcccttccacctgttaaaacatcccaaggttc 238
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 385 TGAOCCTTACATGGGAAATGAATGATGTATGGCATCTAATACTTATGGAAGTAACC 444
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 239 cccttaattcttccaggtatccatctctgggctctgtttgttcaatattcttgag 298
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 445 CTATTTCCGTATATGCGGTACACAAATTGGGGGTGAGTAATGTATTATTTCAGGAT 504
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 299 ccctcccaattgcagtgaaaagaacacacagaactctgataataatgaagccgaataa 358
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
505 CCTTGTCATTTGCGAGAGGAATTAGAACTACAAATAAGSCCTGTCCGAGTAGTCTAGGAA 564
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 359 tgaactctctagtgccctgaaggaatagcttgaatcattctcctcacaatttggttca 418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 565 TGAATATCACACAGCTGTACTGCTGCATCAGGATCTTAATCAACACATTTAGCTTGG 624
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 419 tccctagat-----caaaactacatttggttattctcaccaaaatagtcagtgaag 472
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 625 CGTTTATTCATTCACATCCCTTACTGTAACATAATGSCAACCTCAATAATATGTGATG 684
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 473 ctgttactgcctgtctcttggaatttgattacattgacttgcagcatattgaat 532
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 685 GGACTATGTCCATCTATATGGGTCTGATGACATGATGCTCTCTTAAGTGTGCTGGAAT 744
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 533 tattcaattctctgccttctcacaatttggggtg 566
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 745 TCTGCATTGCTGTGCCCTCTGCTGCTTTGGATG 778
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: October 7, 2001, 04:07:37
Job time: 5662 sec

THIS PAGE BLANK (USPTO)


```
PS Claim 1; SEQ ID 3078; 71pp + CD-ROM; English.
```

```
xx The present sequence is one of a large number of 5' ESTs derived from
```

```
CC mRNAs encoding secreted proteins. An ORF has been identified within the
```

```
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
```

```
CC derived from 30 different tissues. EST sequences usually correspond
```

```
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
```

```
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
```

```
CC well suited for isolating cDNA sequences derived from the 5' ends of
```

```
CC mRNAs and even in those cases where longer cDNA sequences have been
```

```
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
```

```
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
```

```
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
```

```
CC gene therapy and chromosome mapping procedures. They are used to obtain
```

```
CC upstream regulatory sequences and to design expression and secretion
```

```
CC vectors.
```

```
SQ Sequence 372 BP; 105 A; 87 C; 53 G; 117 T; 10 other;
```

```
XX
```

```
XX Query Match 46.6%; Score 280.8; DB 21; Length 372;
```

```
XX Best Local Similarity 96.3%; Pred. No. 3e-66;
```

```
XX Matches 289; Conservative 6; Mismatches 4; Indels 1; Gaps 1.
```

```
OY 1 atgagttcaagaccgcacacagctccgggtgttcttgatattcctccccaagaatcactgct 60
```

```
Db 73 atggattcaagaccgcacacagctccgggtgttcttgatattcctccccaagaatcactgct 132
```

```
OY 61 tcagatatgaatccacagaaccttcagcacgacctttcaactcaagaagcccttgc-a 119
```

```
Db 133 tcaratatgatgccacagaaccttcagcacgacctttcaactcaagaagcccttgc-a 192
```

```
OY 120 aaaattattgtctagaaaaatgaaaatcttagagactaccagatcgtttggaattat 179
```

```
Db 193 aaattattgtctagaaaaatgaaaatcttagagactaccagawccgtttggaattat 252
```

```
OY 180 gacccttcttttgtagtatcttctcttcacctgttaaaacatatccaagtgtcc 239
```

```
Db 253 gnccttctcttttgtagtatcttctcttmacctgtlndaaacatatccaaggttcc 312
```

```
OY 240 ctctattctcttagagatacatctctcgaggcctgttgttcattaatctcgagc 299
```

```
Db 313 ctctattctcttanaratacatcatctcgaggcctgttgttcattaatctcgagc 372
```

```
RESULT 2
```

```
AAF58252 standard; DNA; 936 BP.
```

```
AC AAF58252;
```

```
XX XX
```

```
DT 24-APR-2001 (first entry)
```

```
XX XX
```

```
DE Oligonucleotide D1835.
```

```
KX KW Electron-transfer group; ETM; mismatch; genotyping;
```

```
XX KM gene expression; ss.
```

```
OS Synthetic.
```

```
XX OS
```

```
PN WO200107665-A2.
```

```
XX PD
```

```
PD 01-FEB-2001.
```

```
PF 26-JUL-2000; 2000WO-US20476.
```

```
XX XX
```

```
PR 26-JUL-1999; 99US-0145695.
```

```
XX PR
```

```
PA 17-MAR-2000; 2000US-0190259.
```

```
XX PA
```

```
(CLIN-) CLINICAL MICRO SENSORS INC.
```

```
XX TX
```

```
TX timek RM;
```

```
XX
```

```

DR      WPI: 2001-159728/16.
XX XX   Nucleic acids containing electron-transfer group, useful as labels in
PT PT   hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT PT   a single surface -
XX XX
PS PS     Example 6; Page 127; 159pp; English.
XX XX
CC CC     The present invention relates to a composition comprising two nucleic
CC CC     acids each containing an electron-transfer group (ETM) having
CC CC     different redox potentials. The invention is used for electronic
CC CC     detection of nucleic acids, especially of substitutions (mismatches)
CC CC     and single-nucleotide polymorphisms, e.g. for genotyping,
CC CC     monitoring gene expression.
XX XX
SQ SQ     Sequence 936 BP; 4 A; 139 C; 10 G; 7 T; 776 other;

Query Match          15.6%; Score 94.2; DB 22; Length 936;
Best Local Similarity 1.1%; Pred. No. 5,2e-16;
Matches    6; Conservative 356; Mismatches 209; Indels    0; Gaps    0;

OY       31 ttctgcgtattccccccgaatacctcgcttaagatatgagtcacagaacttcagcc 90
DB        151 ..... : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY       91 acgacctttccaactcaaaaggcccttgcaaaaatatttgtctagaanaaatgaaactcta 150
DB        211 : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY       151 gggactaccagatccctgttggaattagaccttctttagagtatcctccttc 210
DB        271 ..... : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY       211 acctgttaaaccatccaaggttccocttatattcttcaagatatccattctgy 270
DB        331 ..... : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY       271 ggctcgttttgttcattaatctbgagccttcctaattgtagtgaagaaccaca 330
DB        391 wgctla..... : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY       331 gaacctgtataattttagccgaataatgaatcttcttagtgccttgaggaatagct 390
DB        451 ..... : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY       391 ggaatcatctcccctaacttggttcaccagaaccaactaacatttggttatct 450
DB        511 ..... : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY       451 caccaaatagtcagtgtaagcgcttaactgcccgttcttgggaattgttacatg 510
DB        571 ..... : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY       511 atgacttcagcatatgaattacattctcctgccttctcaatttggggtgcac 570
DB        631 ..... : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY       571 tcagagattgtgatgtgacaacagtgtc 601
DB        691 ..... : : : : : : : : : : : : : : : : : : : : : : : : : : : :
                                ..... 721

RESULT      3
ID ID      AAFS8254
XX AC      AAFS8254 standard; DNA; 936 BP.
XX AC      AAFS8254;
DT DT      24-APR-2001 (first entry)
XX DE      Oligonucleotide D1875.
XX

```

KM Electron-transfer group; ETM; mismatch; genotyping;
 KM gene expression; ss.
 OS Synthetic.
 PN WO200107665-A2.
 PD 01-FEB-2001.
 XX
 XX 26-JUL-2000; 2000WO-US20476.
 XX
 XX 26-JUL-1999; 99US-0145695.
 PR 17-MAR-2000; 2000US-0190259.
 XX
 PA (CLIN-) CLINICAL MICRO SENSORS INC.
 XX
 PI Uneek RM;
 XX
 DR WPI; 2001-159728/16.
 XX
 PT Nucleic acids containing electron-transfer group, useful as labels in hybridization assays, e.g. for genotyping, allowing repeat analyses on a single surface
 XX
 PS Example 6; Page 127; 159pp; English.
 XX
 CC The present invention relates to a composition comprising two nucleic acids each containing an electron-transfer group (ETM) having different redox potentials. The invention is used for electronic detection of nucleic acids, especially of substitutions (mismatches) and single-nucleotide polymorphisms, e.g. for genotyping, monitoring gene expression.
 CC
 XX
 SO Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;

Query Match 15.6%; Score 94.2; DB 22; Length 936;
 Best Local Similarity 1.1%; Pred. No. 5.2e-16;
 Matches 6; Conservative 356; Mismatches 209; Indels 0; Gaps 0;
 OY 31 ttctgtattcctccagaatacactgctcagataatgagtcacagacttcagcc 90
 DB 151 www.
 OY 91 acgaccttcaactcaagcccttcgcaaaatattctgtagaataatgaaatccta 150
 DB 211 www.
 OY 151 gggactatccagatcctgttggaatatagaacttcttggaagtatcctcttc 210
 DB 271 www.
 OY 211 acctgttaaacatatacaggttcccttataattcttcagagataccattctg 270
 DB 331 www.
 OY 271 ggcctgtttgttcataattcggagccttcctaatctgacgtgaaagaaacacaca 330
 DB 391 wgctta.
 OY 331 gaacctgataatataagcgaataatgaatcttctagtcgacctgagagacatagct 390
 DB 451 www.
 OY 391 ggaactatcctcactacattggttcaatcctgatacaaaactactgttgattct 450
 DB 511 www.
 OY 451 caccaaaatagtcagtgtaagcgttactgctgtctgtcttggaatttgattacatg 510
 DB 571 www.
 OY 511 atgacttcagcattatgataatcattctctgccttcctcaatttggtggtgcac 570

DB 631 www. 690
 OY 571 tcagagattgtgattgtgacaaatgttgt 601
 DB 691 www. 721
 RESULT 4
 ID AAF58257 standard; DNA; 936 BP.
 XX
 AC AAF58257;
 XX
 DT 24-APR-2001 (first entry)
 XX
 DE Oligonucleotide D1954.
 XX
 KM Electron-transfer group; ETM; mismatch; genotyping;
 KM gene expression; ss.
 OS Synthetic.
 PN WO200107665-A2.
 PD 01-FEB-2001.
 XX
 XX 26-JUL-2000; 2000WO-US20476.
 XX
 XX 26-JUL-1999; 99US-0145695.
 PR 17-MAR-2000; 2000US-0190259.
 XX
 PA (CLIN-) CLINICAL MICRO SENSORS INC.
 XX
 PI Uneek RM;
 XX
 DR WPI; 2001-159728/16.
 XX
 PT Nucleic acids containing electron-transfer group, useful as labels in hybridization assays, e.g. for genotyping, allowing repeat analyses on a single surface
 XX
 PS Example 6; Page 127; 159pp; English.
 XX
 CC The present invention relates to a composition comprising two nucleic acids each containing an electron-transfer group (ETM) having different redox potentials. The invention is used for electronic detection of nucleic acids, especially of substitutions (mismatches) and single-nucleotide polymorphisms, e.g. for genotyping, monitoring gene expression.
 CC
 XX
 SO Sequence 936 BP; 5 A; 142 C; 7 G; 6 T; 776 other;
 Query Match 15.6%; Score 94.2; DB 22; Length 936;
 Best Local Similarity 1.1%; Pred. No. 5.2e-16;
 Matches 6; Conservative 356; Mismatches 209; Indels 0; Gaps 0;
 OY 31 ttctgtattcctccagaatacactgctcagataatgagtcacagacttcagcc 90
 DB 151 www.
 OY 91 acgaccttcaactcaagcccttcgcaaaatattctgtagaataatgaaatccta 150
 DB 211 www.
 OY 151 gggactatccagatcctgttggaatatagaacttcttggaagtatcctcttc 210
 DB 271 www.
 OY 271 acctgttaaacatatacaggttcccttataattcttcagagataccattctg 270
 DB 331 www. 390

Qy	271	ggccctgtttgttcattcaattccggagccttcctaattgacgtggaagaanaaccaca	333
Db	331	wgcttaa	450
Qy	331	gaacctcgtataatagagccgaataatgaatcttctagtgccctgagagcaatagct	390
Db	451	aa	510
Qy	331	ggatcaattctcccaatttggttcatcctatgatacaaaactaatgtgtattct	450
Db	511	aa	570
Qy	451	caccaaatagtcagtgtaagcgtctactgcctctgcttggtggaaattgatatcatg	510
Db	571	aa	630
Qy	511	atgaccttcagcattatgaattatcattctctgccttcctaatttgagggtgcac	570
Qy	631	aa	690
Qy	571	tcagaggaattgtgattgtgacaattgttct 601	
Db	691	aaa 721	
RESULT 5			
ID	AAF58259	standard; DNA; 936 BP.	
XX	AAF58259;		
AC			
XX			
DT	24-APR-2001	(first entry)	
XX			
DE	Oligonucleotide D2004.		
XX			
KM	Electron-transfer group; ETM; mismatch; genotyping;		
XX	gene expression; ss.		
OS	Synthetic.		
XX			
PN	W0200107665-A2.		
XX			
PD	01-FEB-2001.		
XX			
PE	26-JUL-2000; 2000MO-US20476.		
XX			
XX	26-JUL-1999; 99US-0145695.		
XX	17-MAR-2000; 2000US-0190259.		
PA	(CLIN-) CLINICAL MICRO SENSORS INC.		
PI	Umek RM;		
XX			
DR	WPI; 2001-159728/16.		
XX			
PT	Nucleic acids containing electron-transfer group, useful as labels in		
PT	hybridization assays, e.g. for genotyping, allowing repeat analyses on		
PS	a single surface		
XX			
XX	Example 6; Page 128; 159pp; English.		
CC	The present invention relates to a composition comprising two nucleic		
CC	acids each containing an electron-transfer group (ETM) having		
CC	different redox potentials. The invention is used for electronic		
CC	detection of nucleic acids, especially of substitutions (mismatches)		
CC	and single-nucleotide polymorphisms, e.g. for genotyping,		
CC	monitoring gene expression.		
XX			
XX	Sequence 936 BP; 6 A; 138 C; 8 G; 8 T; 776 other;		
XX			
Query Match	15.6%;	Score 94.2;	DB 22; Length 936;

OY	31	tcttcgtgattcccccagaaatcacgcgtccttcagaatatggatccacagacttcaaggcc	90
Db	151	www.....	210
OY	91	acgacctttccaactcaagcccccttgcaaaaattatttgcagttagaaaaatgaactcta	150
Db	211	www.....	270
OY	151	gggactaccagatcctgtttggaataagaccttctttggagtatcttcctttc	210
Db	271	www.....	330
OY	211	acctgttaaaacccataccaaaggttcccttatattcttcacagatalccattctg	270
Db	331	www.....	390
OY	271	ggctctgttttgtcatcaatctcgagccttcctaattgcagtgaaagaaacacaca	330
Db	391	wgctta.....	450
OY	331	gaacctctgataatatgagccgaataatgaatcttcttagtgcctcgagagcaatagct	390
Db	451	www.....	510
OY	391	ggaactcttcctccatcttggtttccatccagatacaaaactaacattgtgattctc	450
Db	511	www.....	570
OY	451	caccaaatagtcagtgtaagcgtgttacgcbctgtctcttggaatttgattacattg	510
Db	571	www.....	630
OY	511	atgacttcagcatattgaattatcaattctcgccttcctaatttgggtgccac	570
Db	631	www.....	690
OY	571	tcagagattgtgatgtgcaacaatgtgtt	601
Db	691	www.....	721

RESULT 6
AAFS8262
ID AAF58262 standard; DNA; 936 BP.

AC AAF58262;
XX
DT 24-APR-2001 (first entry)
XX
DE Oligonucleotide D2007.
XX
KW Electron-transfer group; ETM; mismatch; genotyping;
gene expression; ss.
XX
CS Synthetic.
XX
PN WO200107665-A2.
PD
PD 01-FEB-2001.
PF 26-JUL-2000; 2000WO-US20476.
PR 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
PA (CLIN-) CLINICAL MICRO SENSORS INC.
PI Umek RM;
XX WPI; 2001-159728/16.
OR

[illegible]

Db 631 www. 690
Oy 571 tcagagatgtgattgtgacaatgtgtt 601
Db 691 www. 721

RESULT 8
AAF58252/c
ID AAF58252 standard; DNA; 936 BP.

XX AAF58252;

XX 24-APR-2001 (first entry)

XX Oligonucleotide D1835.

XX Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.

XX Synthetic.

XX WO200107665-A2.

XX 01-FEB-2001.

XX 26-JUL-2000; 2000MO-US20476.

XX 26-JUL-1999; 99US-0145695.

XX 17-MAR-2000; 2000US-0190259.

XX (CLIN-) CLINICAL MICRO SENSORS INC.

XX Umek RM;

XX WPI; 2001-159728/16.

XX Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface -

XX Example 6; Page 127; 159pp; English.

XX The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
and single-nucleotide polymorphisms, e.g. for genotyping,
monitoring gene expression.

SQ Sequence 936 BP; 4 A; 139 C; 10 G; 7 T; 776 other;

Query Match 15.5%; Score 93.4; DB 22; Length 936;

Best Local Similarity 0.7%; Pred. No. 8.5e-16; Mismatches 209; Indels 0; Gaps 0;

Matches 4; Conservative 358; Mismatches 209; Indels 0; Gaps 0;

Oy 31 ttctgtattcttcgcagaatcactgcttcagaatagtcacagaacttcagcc 90

Db 675 www. 616

Oy 91 acgaccttcaactcaagcccttcgcaaaattattgctagaanaaaatgaaatctta 150

Db 615 www. 556

Oy 151 gggacttcagatccgttggaatagatgacttcttcttgagttatcttccttctc 210

Db 555 www. 496

Oy 211 acctgttaaacatatccaggttcccttatattcttcagatatcatcttcg 270

Db 495 www. 436

Oy 271 ggcctgttttctaatcctcgagcccttcctaatctgcagtgaagaacaccaca 330
Db 435 www. 376
Oy 331 gaactcgtataatctgagccgagataatgacttctcttagtgcctcgagagcaatgct 390
Db 375 www. 316
Oy 391 ggaatcctccctcacatttggttccctagatacaaacatttgytattct 450
Db 315 www. 256
Oy 451 caccacaatagtcagtglaagcgtctacgtctccttcttggaatttgatcata 510
Db 255 www. 196
Oy 511 atgacttcagcatattgaattcattctctgccttcatttgagggtccac 570
Db 195 www. 136
Oy 571 tcagagatgtgattgtgacaatgtgtt 601
Db 135 www. 105

RESULT 9

AAF58254/c
ID AAF58254 standard; DNA; 936 BP.

XX AAF58254;

XX 24-APR-2001 (first entry)

XX Oligonucleotide D1875.

XX Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.

XX Synthetic.

XX WO200107665-A2.

XX 01-FEB-2001.

XX 26-JUL-2000; 2000MO-US20476.

XX 26-JUL-1999; 99US-0145695.

XX 17-MAR-2000; 2000US-0190259.

XX (CLIN-) CLINICAL MICRO SENSORS INC.

XX Umek RM;

XX WPI; 2001-159728/16.

XX Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface -

XX Example 6; Page 127; 159pp; English.

XX The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
monitoring gene expression.

SQ Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;

Query Match 15.5%; Score 93.4; DB 22; Length 936;

Best Local Similarity 0.7%; Pred. No. 8.5e-16;

```

PT Nucleotides containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface.
XX
PS
XX Example 6; Page 127; 159pp; English.
XX
CC The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
CC
XX
XX Sequence 936 BP; 5 A; 142 C; 7 G; 6 T; 776 other;
SO

Query Match 15.5%; Score 93.4; DB 22; Length 936;
Best Local Similarity 0.7%; Pred. No. 8.5e-16;
Matches 4; Conservative 358; Mismatches 209; Indels 0; Gaps 0;

QY 31 ttctcgtatttcctccagaatactcgtctcagaatatgatccacagaacttcaagc 90
Db 675 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 616
QY 91 acgaccttccaactcaagcccttcgaaataatttctagaaaaatgaactcta 150
Db 615 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 556
QY 151 gggactaccagactcgttttggaatatagccttcttcttgtagtattcctcttc 210
Db 555 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 496
QY 211 acctgttaaacacataccagaagttccctctatattcttcagatccattctcg 270
Db 495 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 436
QY 271 ggcctgtttgttcataatctcgaagccttcctaattgcagtgaagaacaacca 330
Db 435 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 376
QY 331 gaaactcgtataattgagccgaataatgaactcttcttagtgcctcgagacaatagc 390
Db 375 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 316
QY 391 ggaactatctccccaacttggttcaactcctagatacgaactacatttgttatct 450
Db 315 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 256
QY 451 caccaaatagtcagtgtaagcgtctactcctcgttcttcgggaatttattacatg 510
Db 255 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 196
QY 511 atgacttcagcatattgaattatcattctcctgccttccaatttgggggtgccac 570
Db 195 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 136
QY 571 tcagagatgtgatttgtaacaatgttgtt 601
Db 135 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 105

RESULT 11
ID AAF58259/c
XX AAF58259;
XX
XX 24-APR-2001 (first entry)
XX
XX Oligonucleotide D2004.
XX
XX Electron-transfer group; ETM; mismatch; genotyping;
XX gene expression; ss.
XX

```



```
Db 435 WWWWWW... 376
Oy 331 gaaacttgataatgagccgataatagaactcttctgtgcccgaagcagatgct 390
Db 375 WWWWWW... 316
Oy 391 ggaatcctctccacattgttctatccatagatacaaacacacattgtgtattct 450
Db 315 WWWWWW... 256
Oy 451 caccaaaatagtcagtgtaagcgtgtactgtctgtcttggaatttgatcatg 510
Db 255 WWWWWW... 196
Oy 511 atgaacttcgacattatgaataatcattctctgccttcacatttgagggtgcac 570
Db 195 WWWWWW... 136
Oy 571 tcagagagattgtatgtgacaaatgttgt 601
135 WWWWWW... 105
```

RESULT 13

AAFS8255/c
ID AAF58255 standard; DNA; 938 BP.

AAFS8255;

24-APR-2001 (first entry)

Oligonucleotide D1876.

Electron-transfer group; ETM; mismatch; genotyping;
gene expression; ss.

Synthetic.

WO200107665-A2.

01-FEB-2001.

26-JUL-2000; 2000WO-US20476.

26-JUL-1999; 99US-0145695.

17-MAR-2000; 2000US-0190259.

(CLIN-) CLINICAL MICRO SENSORS INC.

Umek RM;

WPI; 2001-159728/16.

Nucleic acids containing electron-transfer group, useful as labels in hybridization assays, e.g. for genotyping, allowing repeat analyses on a single surface

Example 6; Page 127; 159pp; English.

The present invention relates to a composition comprising two nucleic acids each containing an electron-transfer group (ETM) having different redox potentials. The invention is used for electronic detection of nucleic acids, especially of substitutions (mismatches) and single-nucleotide polymorphisms, e.g. for genotyping, monitoring gene expression.

Sequence 938 BP; 4 A; 144 C; 9 G; 5 T; 776 other;

Query Match

Best Local Similarity 0.78; Pired. No. 8.5e-16;
Matches 4; Conservative 358; Mismatches 209; Indels 0; Gaps 0;

```
Oy 31 ttctgtattctccccaatcaactgcttcagaatagtcacagaaattcagcc 90
Db 675 WWWWWW... 616
Oy 91 acgaccttccaactcaagcccttgcaaaatatttctagaaaatgaactta 150
Db 615 WWWWWW... 556
Oy 151 gggactacagacatcgtttggaattagaaccttctcttgagtgatcctcttc 210
Db 555 WWWWWW... 496
Oy 211 acctgttaaacacataccaagttcccttatatttcttcagatatcattctg 270
Db 495 WGW... 436
Oy 271 ggcctgtttgttcaataatctcgagccttcctaattgcaatgaagaaacaca 330
Db 435 WWWWWW... 376
Oy 331 gaaacttgataatgagccgataatagaactcttctgtgcccgaagcagatgct 390
Db 375 WWWWWW... 316
Oy 391 ggaatcctctccacattgttctatccatagatacaaacacacattgtgtattct 450
Db 315 WWWWWW... 256
Oy 451 caccaaaatagtcagtgtaagcgtgtactgtctgtcttggaatttgatcatg 510
Db 255 WWWWWW... 196
Oy 511 atgaacttcgacattatgaataatcattctctgccttcacatttgagggtgcac 570
Db 195 WWWWWW... 136
Oy 571 tcagagagattgtatgtgacaaatgttgt 601
135 WWWWWW... 105
```

RESULT 14

AAAI6693
ID AAAI6693 standard; cDNA; 1330 BP.

AAAI6693;

16-JUN-2000 (first entry)

Human secreted protein clone pe246_1 nucleotide sequence SEQ ID NO:151.

Human; secreted protein; immunostimulant; immunosuppressant; virucide; antibacterial; antifungal; cytostatic; antiinflammatory; dermatological; antidiabetic; antiasthmatic; antiarthritic; antirheumatic; protozoacide; antihypertoid; immune deficiency; severe combined immunodeficiency; SCID; infection; HIV; hepatitis; malaria; autoimmune disorder; systemic lupus; connective tissue disease; multiple sclerosis; erythematosis; rheumatoid arthritis; autoimmune pulmonary inflammation; asthma; Guillain-Barre syndrome; autoimmune thyroiditis; myasthenia gravis; insulin dependent diabetes mellitus; graft-versus-host-disease; autoimmune inflammatory eye disease; allergy; ss.

Homo sapiens.

WO200009552-A1.

24-FEB-2000.

13-AUG-1999; 99WO-US18298.

14-AUG-1998; 98US-0096622.

17-AUG-1998; 98US-0096815.

PR 04-SEP-1998; 98US-00992229.
 PR 23-OCT-1998; 98US-0105368.
 PR 08-JAN-1999; 99US-0115234.
 PR 12-FEB-1999; 99US-0119931.
 PR 18-FEB-1999; 99US-0120575.
 PR 30-APR-1999; 99US-0132020.
 PR 11-AUG-1999; 99US-0096622.
 XX
 PA (GEMV) GENETICS INST INC.
 XX
 PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Mendeig D, Treacy M, Agostino MJ, Steinger RJ, Spaulding V;
 PI Wong GG, Clark HF, Fechtel K;
 XX
 DR MPI: 2000-205979/18.
 DR P-PSDB; AAY94973.
 XX
 PF New polynucleotides encoding secreted proteins, which may have e.g.
 PF nutritional, chemokine, immune stimulating or suppressing,
 PF hematopoiesis regulating, tissue growth, activin/inhibin
 PF antiinflammatory or tumor inhibition activity -
 PS
 PS Claim 160; Page 606-607; 641pp; English.
 XX
 CC AA16618 to AA16697 encode the human secreted proteins given in
 CC AAY9498 to AAY94980, isolated from human adult brain, adult thyroid,
 CC adult retina, foetal carcinoma, adult blood, adult neural, foetal kidney,
 CC adult placenta, adult testis, whole embryo, adult cartilage, kidney,
 CC foetal brain, adult thymus, foetal placenta, adult uterus, adult tumour,
 CC and adult bladder, cDNA libraries. The polynucleotides and proteins are
 CC predicted to have biological activities which would make them suitable
 CC for treating, preventing or ameliorating medical conditions in humans
 CC and animals. The polynucleotides can be used as markers for tissues in
 CC which the protein is preferentially expressed, as molecular weight
 CC markers on Southern gels, and as chromosome markers or tags to identify
 CC chromosomes or to map gene positions. The proteins can be used in the
 CC treatment of immune deficiencies and disorders, such as severe combined
 CC immunodeficiency (SCID), as well as viral, bacterial, fungal and other
 CC infections. These infections include human immunodeficiency virus (HIV),
 CC hepatitis, herpesviruses, mycobacteria, Leishmania spp., malaria and
 CC candidiasis. The proteins can be used to treat autoimmune disorders such
 CC as connective tissue disease, multiple sclerosis, systemic lupus
 CC erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation,
 CC Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent
 CC diabetes mellitus, myasthenia gravis, graft-versus-host-disease and
 CC autoimmune inflammatory eye disease. The proteins can also be used to
 CC treat allergic conditions, such as asthma. AA16698 to AA16774 represent
 CC probes for the human secreted proteins from the present invention.
 SO
 SO Sequence 1330 BP; 364 A; 285 C; 309 G; 372 T; 0 other:
 Query Match 14.3%; Score 86.4; DB 21; Length 1330;
 Best Local Similarity 51.1%; Pred. No. 7, 2e-14;
 Matches 232; Conservative 0; Mismatches 216; Indels 6; Gaps 1
 DY 119 aaaaatattgtcagaanaaataatccttagagactacacagacccgtgttgaata 178
 DB 485 agaagcttcgaagagagacacccaagtccttgagggttgtagagactcgtacgtcc 544
 OY 179 tgaactttcttttgagtgatctctcctttcaacttgtaaacacatatacaagtttc 238
 DB 545 tgaagcttagacatggaataaacaatgagtgtatgtagacatcaatacattagaaagtaacc 604
 OY 239 ccttatattcttcacagatatccatctcgtgggcctcttggtgataataattcggag 298
 DB 605 ctattccgtgtatatacgtgtacacaaatttgggggtcagtaattgtattatttcaggat 664
 OY 299 ccttcctaatgtgcagtgaagaagaaacacacagaacactctgataatatgtgacgcgaata 358
 DB 665 cctgttcacatgtgcacaggaatttagaacatacaaaagcctgtgtccgaggttagtctaggaa 724
 DY 359 tgaactctctatgtccctcgaagacataacgtcggagatcatcttcctccgatttgcattca 418

Db	725	tgaatcacacagctcgtcgtacgtgcacacaggtatcctaacaacattagcttgg	784
Oy	419	tcctagat-----cgaactacattgtyggtatcttcacacaatagtcagtgaag	472
Db	785	cgtttattcattccacacccttactgactactatgacacgaataatgtcacy	844
Oy	473	ctgtgactgcctgtcttgggaatttgattcattcagatgacttcacatattgaat	532
Db	845	ggactatgctcatcttaatgagtgctgtagtgcattgtgtctcccttaagtgcgtgaat	904
Oy	533	tattcaattcctgcctcttctcaatttggagtg	566
Db	905	tctgattgctgtgtccctctctgccttggatg	938
RESULT 15			
XX	AA15104	standard; cDNA; 1669 BP.	
XX	AA15104		
AC	AA15104;		
XX	19-APR-1999	(first entry)	
XX	19-APR-1999		
DE	High affinity immunoglobulin E receptor-like protein (IGERB) cDNA.		
XX	High affinity immunoglobulin E receptor-like protein; IGERB;		
KW	Incye clone 927955; inflammatory response; AIDS; Addisons's disease;		
KW	atherosclerosis; bronchitis; ulcerative colitis; diabetes mellitus;		
KW	emphysema; gout; Graves's disease; osteoporosis; rheumatoid arthritis		
KW	Sjogren's syndrome; cancer; ss.		
XX	Homio sapiens.		
OS	Homio sapiens.		
XX	Key	Location/Qualifiers	
FM	159..878	/*tag= a	
FT	CDS	/product= IGERB	
XX	US5871930-A.		
XX	16-FEB-1999.		
PD	21-AUG-1997;	97US-0916902.	
XX	21-AUG-1997;	97US-0916902.	
XX	(INCY-) INCYTE PHARM INC.		
PA	Bandman O, Corley NC, Lal P;		
PI	WPI: 1999-16628/14.		
XX	P-PSDB; AAW96745.		
DR	New polynucleotide coding for high affinity immunoglobulin E		
PT	receptor-like protein - and its complement, useful for diagnosis,		
PT	prevention and antisense therapy of inflammatory responses		
XX	Claim 4; Fig 1A-D; 30pp; English.		
PS	The present sequence encodes a high affinity immunoglobulin E		
XX	receptor-like protein (IGERB). The present sequence was first		
CC	identified in incye clone 927955 for the brain cDNA library		
CC	BRINOT04. Polynucleotides complementary to the IGERB cDNA can		
CC	be used as probes to IGERB gene expression. IGERB proteins and		
CC	nucleotides may be used for the diagnosis of inflammatory		
CC	responses associated with expression of IGERB, e.g AIDS,		
CC	Addison's disease, atherosclerosis, bronchitis, ulcerative colitis,		
CC	diabetes mellitus, emphysema, gout, Graves's disease, osteoporosis,		
CC	rheumatoid arthritis, Sjogren's syndrome and complications of cancer.		
CC	The sequences may also be useful in assays that detect activation or		
CC	induction of various cancers.		

Sequence 1669 BP; 481 A; 350 C; 334 G; 501 T; 3 other;

Query Match 14.3%; Score 86.4; DB 20; Length 1669;
Best Local Similarity 51.1%; Pred. No. 7.8e-14;
Matches 232; Conservative 0; Mismatches 216; Indels 6; Gaps 1;

```
QY 119 aaaaattattgtagaanaatgaanaatcttaggactaccagatcctgttggaaat 178
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 325 agaattcttgaaggagaaacccaagctctggggttggagatctgacgcctga 384
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 179 tgaccttctcttggagttatctctcttccacctgttaaaacalaccaagttc 238
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 385 tgaaccttagcatggggaataacaatgatgtatgcatctaatactatggaatacc 444
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 239 ccttataattcttcaaggatcaccatctcgggctcgtttgttcataattctggag 298
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 445 ctatttcggtatatacgggtacacaatttgggggtcagtaacgttatatttcaggat 504
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 299 ccttcctaattgagtgaaanaaaacacagaaactcgtataatttgagccgaata 358
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
    505 ccttgcgaattgacagcagaataatagaactacaaaaggctgtccgaggtagctagaa 564
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 359 tgaatcttcttagtgccctgaagagaatagcttgaatcaltctcctcacatttggttca 418
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 565 tgaatatcaaccagctctgtactgctgcacatcaggatcttaatcaacacatttagcttg 624
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 419 tccctagat-----caaaactacatttgggtatctcaccaaaatagtcagtgaaag 472
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 625 cgtttatcattccaccacccttactgttaactactatgccaactcaataatgtlcatg 684
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 473 ctgttactgctcgttcttgggaatttggattacatgagacttccagcattatgaat 532
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 685 ggactatgttccatcttaaggggtcgtgatgtgccttgccttaagtgtgtcgtgaat 744
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 533 tattcaattctctgccttctcacaatttgggggtg 566
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 745 tctcgatgtgtctccctctctgccttcttgatg 778
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: October 7, 2001, 04:07:30
Job time: 2810 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 7, 2001, 03:19:45 ; Search time 37.7 Seconds
(without alignments)
3027.974 Million cell updates/sec

Title: US-09-735-712-1

Perfect score: 603

Sequence: 1 atgattcaagcaccgcaca.....attgtgaacatgtttga 603

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/prodata/2/lna/5A.COMB.seq:*
2: /cgn2_6/prodata/2/lna/5B.COMB.seq:*
3: /cgn2_6/prodata/2/lna/6A.COMB.seq:*
4: /cgn2_6/prodata/2/lna/6B.COMB.seq:*
5: /cgn2_6/prodata/2/lna/FCIUS.COMB.seq:*
6: /cgn2_6/prodata/2/lna/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86.4	14.3	1669	2	US-08-916-902A-2
2	86.4	14.3	1669	2	US-09-213-389-2
3	63.2	10.5	1661	1	US-08-318-492-3
4	63.2	10.5	1661	1	US-08-707-340-3
5	63.2	10.5	1661	2	US-08-994-578-3
6	49.2	8.2	2545	1	US-07-869-933-22
7	49.2	8.2	2545	4	US-09-103-663-22
8	42.6	7.1	7218	1	US-08-232-463-14
9	34.6	5.7	246240	2	US-08-724-394A-20
10	34.6	5.7	246240	2	US-08-724-394A-21
11	34.6	5.7	246240	2	US-08-724-394A-22
12	33.6	5.6	1711	1	US-08-568-147B-1
13	33.4	5.5	6669	3	US-09-212-971-5
14	33.4	5.5	6669	4	US-08-800-929A-5
15	32.2	5.3	3947	4	US-08-975-762-47
16	32.2	5.3	8920	4	US-08-446-855A-1
17	32	5.3	9793	1	US-09-150-741-1
18	32	5.3	9793	1	US-08-470-202-56
19	32	5.3	9793	1	US-08-471-770-56
20	32	5.3	9793	2	US-08-468-059-56
21	31.4	5.2	3393	4	US-09-104-324B-1
22	31.4	5.2	3393	4	US-09-162-713-1
23	31.4	5.2	3580	4	US-09-081-345-1
24	31.4	5.2	4291	2	US-08-417-210A-81
25	31.4	5.2	4517	4	US-09-140-804-9
26	31	5.1	1902	4	US-08-793-044-4
27	31	5.1	2839	4	US-09-061-702-1

28	31	5.1	2945	3	US-09-058-489-85	Sequence 85, Appl
29	31	5.1	9439	3	US-09-058-489-89	Sequence 89, Appl
30	30.8	5.1	699	5	PCT-US95-13335-2	Sequence 2, Appl
31	30.8	5.1	2662	3	US-08-750-357-8	Sequence 8, Appl
32	30.8	5.1	3466	2	US-08-468-036-38	Sequence 38, Appl
33	30.6	5.1	4526	1	US-07-855-412B-4	Sequence 4, Appl
34	30.6	5.1	4526	2	US-08-308-887A-4	Sequence 4, Appl
35	30.6	5.1	4526	3	US-08-881-094-4	Sequence 4, Appl
36	30.4	5.0	2189	3	US-08-846-020A-1	Sequence 1, Appl
37	30.4	5.0	3427	3	US-09-009-913-2	Sequence 2, Appl
38	30.4	5.0	5510	3	US-09-009-913-3	Sequence 3, Appl
39	30.4	5.0	5667	3	US-09-009-913-4	Sequence 4, Appl
40	30.4	5.0	22846	2	US-08-469-461-3	Sequence 3, Appl
41	30.4	5.0	22846	3	US-07-890-609-3	Sequence 3, Appl
42	30.2	5.0	315	2	US-08-743-200-1	Sequence 1, Appl
43	30.2	5.0	376	2	US-08-623-906A-18	Sequence 18, Appl
44	30.2	5.0	1956	1	US-08-318-831-7	Sequence 7, Appl
45	30	5.0				

ALIGNMENTS

RESULT 1
US-08-916-902A-2
; Sequence 2, Application US/08916902A
; Patent No. 5871930
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
; TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/916,902A
; FILING DATE: Herewith
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0371 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELETYPE:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1669 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT04
; CLONE: 927955
; US-08-916-902A-2

```

Query Match      14.3%; Score 86.4; DB 2; Length 1669;
Best Local Similarity 51.1%; Pred. No. 5.9e-15;
Matches 232; Conservative 0; Mismatches 216; Indels 6; Gaps 1;

QY 119 aaaaattattgctagaanaaataaaatccttaggactaccagatccgtttggaatta 178
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 325 AGAAGTTCCTTGAAGGGAGAACCCAAAGTCTTGGGGTTGTCAGATTCTAGCCCTGA 384
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 179 tgaacctttcttggagttatctctcttcccttgaacctgttaaacacatccagagtttc 238
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 385 TGAGCCTTAGCATGGGAATACATATGTGTATGGCATTAATTAATTAAGAGTAACC 444
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 239 ccttatattcttccagagatccatctcctgggctcgtgtttgttcattatctgag 298
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 445 CTATTTCGGGTATATGCGGTACACAAATTGGGGGTGAGTAAGTTATTAATTTACAGAT 504
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 299 ccttctaatgtcagtgaaagaaacacacagaaactctgataatattgagccgaataa 358
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    505 CCTTGTCAATTGACGAGCAAGAAATTAAGAACTCAAAAGGCCGTCCGAGGTAGCTAGGAA 564
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    359 tgaatcttcttagtgccctgagagcaatagcttggaatcattctccatcatttggttca 418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 565 TGAATATCACCACTCTGTACTGCTGCATCAGGATCTTAATACACATTTAGCTTG 624
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 419 tccatagat-----caaaactacattgtgttattctccacaaatagtcagtgtaag 472
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 625 CGTTTATTCATTCCATCCACCCTTACTGTAACTAATGCAACTCAAAATAATTTGCATG 684
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 473 ctgttactgctcgttcttgggaatttgaattgaattgaacttccagattatgaat 532
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 685 GGACTATGTCATCTTAATGGGTCTGTGATGGCATGGTCTCTCTTAAGTGTCTGGAAT 744
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 533 tatcatctctcgtccttctccaatttggggtc 566
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 745 TCTGCAATGCTGTGCTCCTCTCTGCTTGTGATG 778
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 2
US-09-213-389-2
; Sequence 2, Application US/09213389
; Patent No. 5977072
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
; TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Inocyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/213,389
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/916,902
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0371 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555

```

```

TELEFAX: 415-845-4166
;
; TELEEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1669 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT04
; CLONE: 927955
;
US-09-213-389-2

Query Match      14.3%; Score 86.4; DB 2; Length 1669;
Best Local Similarity 51.1%; Pred. No. 5.9e-15;
Matches 232; Conservative 0; Mismatches 216; Indels 6; Gaps 1;

QY 119 aaaaattattgctagaanaaataaaatccttaggactaccagatccgtttggaatta 178
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 325 AGAAGTTCCTTGAAGGGAGAACCCAAAGTCTTGGGGTTGTCAGATTCTAGCCCTGA 384
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 179 tgaacctttcttggagttatctctcttcccttgaacctgttaaacacatccagagtttc 238
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 385 TGAGCCTTAGCATGGGAATACATATGTGTATGGCATTAATTAATTAAGAGTAACC 444
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 239 ccttatattcttccagagatccatctcctgggctcgtgtttgttcattatctgag 298
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 445 CTATTTCGGGTATATGCGGTACACAAATTGGGGGTGAGTAAGTTATTAATTTACAGAT 504
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 299 ccttctaatgtcagtgaaagaaacacacagaaactctgataatattgagccgaataa 358
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    505 CCTTGTCAATTGACGAGCAAGAAATTAAGAACTCAAAAGGCCGTCCGAGGTAGCTAGGAA 564
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    359 tgaatcttcttagtgccctgagagcaatagcttggaatcattctccatcatttggttca 418
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 565 TGAATATCACCACTCTGTACTGCTGCATCAGGATCTTAATACACATTTAGCTTG 624
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 419 tccatagat-----caaaactacattgtgttattctccacaaatagtcagtgtaag 472
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 625 CGTTTATTCATTCCATCCACCCTTACTGTAACTAATGCAACTCAAAATAATTTGCATG 684
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 473 ctgttactgctcgttcttgggaatttgaattgaattgaacttccagattatgaat 532
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 685 GGACTATGTCATCTTAATGGGTCTGTGATGGCATGGTCTCTCTTAAGTGTCTGGAAT 744
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 533 tatcatctctcgtccttctccaatttggggtc 566
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 745 TCTGCAATGCTGTGCTCCTCTCTGCTTGTGATG 778
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
US-08-318-492-3
; Sequence 3, Application US/08318492
; Patent No. 5552312
; GENERAL INFORMATION:
; APPLICANT: Lim, Bing
; APPLICANT: Adra, Chaker N.
; APPLICANT: Lelias, Jean-Michel
; TITLE OF INVENTION: RECOMBINANT HTM4 GENE, PROTEIN AND
; TITLE OF INVENTION: ASSAYS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P. C.
; STREET: Two Millitia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```



```

QY 6 tcaagacgcacagtcgcgtgttctgtatctccacgaatacactgcttcaaga 65
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 875 TTAAGATTTCATACACAGCAGTAGTTTACACCATTTACATTAAGATAAGA 934
QY 66 atagatccacagacttcagccagacttcaactgaagccctcgcaaaatt 125
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 935 TTCAAAATATTATTAAACGGTTTACGTTGAATGTCCCATGAGTCGGCTACTAAT 994
QY 126 attgtcagaaataaataccttagagactatccagatcctgttctgaatgaacct 185
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 995 ATTTCCTCTGTTGCATACGCTCACAGATTAATTCGAGCTTGCTGCAGTGCAG 1054
QY 186 tctcttggaatctctccttctcactgtttaaaccatatacgaagttcccttat 245
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1055 GAGACTGCGATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1114
QY 246 attcttcagagatcatctcgtggtctgttcttctcaatctcgtagactcct 305
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1115 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1174
QY 306 aatgcagtgaaagaaacacagaaactcgtataataltgagccgaataatgaact 365
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1175 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1234
QY 366 tcttagtgcctgagagcaatagctgaatcatctcctcactggttcaactaga 425
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1235 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1294
QY 426 tcaaacactacattgttgtaattctccacaaatagtcagtgtaagctgtactgcct 485
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1295 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1354
QY 486 gtcttggaatttgatcatcatgactgactcagcaattatgaattatcattcct 545
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1355 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1414
QY 546 gccttctcaattt 560
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1415 YYYYYYYYYYYYYY 1429

```

```

NAME: Fitts, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 017957-000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 246240 base pairs
TYPE: nucleic acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1..246240
OTHER INFORMATION: /note= "HLA-H.CONFIG"
US-08-724-394A-20

Query Match          5.7%; Score 34.6; DB 2; Length 246240;
Best Local Similarity 55.4%; Pred. No. 6.8;
Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 449 ctcaacaaatagtcagtgtaaggctgtactgtcctgtcttcttggaatttgatcat 508
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36690 CTCCTCAGATCTCGTATTATTAATTCATTTATGCTGTGTAGTAATTAACCAACTG 36749
QY 509 tgatgaattcagcatatgaattatcatctcctgccttctcaatttggggcgc 568
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36750 TGATGGCTGAACACACAAATAATTTATTATTACAACTTAAGCTCAGAACTCC 36809
QY 569 a 569
Db 36810 A 36810

RESULT 10
US-08-724-394A-21
; Sequence 21, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Krommal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchinashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:

```

RESULT 12
 US-08-568-147B-1/C
 Sequence 1, Application US/08568147B
 Patent No. 5783422
 GENERAL INFORMATION:
 APPLICANT: Sumitani, Yoshihori
 APPLICANT: Kato, Hiroshi
 APPLICANT: Sekiguchi, Kiyoshi
 APPLICANT: Takeda, Katsumichi
 TITLE OF INVENTION: DNA FRAGMENT CODING FOR SCUMOUS CELL
 NUMBER OF SEQUENCES: 17
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Scully, Scott, Murphy & Presser
 STREET: 400 Garden City Plaza
 CITY: Garden City
 STATE: New York
 COUNTRY: USA
 ZIP: 11530
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/568,147B
 FILING DATE:
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 800, 952
 FILING DATE: 02-DEC-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Digiglio, Frank S.
 REGISTRATION NUMBER: 31,346
 REFERENCE/DOCKET NUMBER: 8425
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 516-742-4343
 TELEFAX: 516-742-4366
 TELEX: 230 901 SANS UR
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1711 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: both

Db 2886 ATGTGAGTATTAACCTATACCTCTGAAACATTTTGAACTTTGTTTCTAA 2827

Job time: 2878 sec

QY 248 ttcttcagga 258

Db 2826 ATGTTCTGAA 2816

RESULT 15

US-08-975-762-47

; Sequence 47, Application US/08975762

; Patent No. 6207169

; GENERAL INFORMATION:

; APPLICANT: Reed, Steven G.

; APPLICANT: Lodes, Michael J.

; APPLICANT: Houghton, Raymond

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND TREATMEN

; NUMBER OF SEQUENCES: 73

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SEED and BERRY LLP

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: Washington

; COUNTRY: USA

; ZIP: 98104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/975,762

; FILING DATE: 21-MAR-1997

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Maki, David J.

; REGISTRATION NUMBER: 31,392

; REFERENCE/DOCKET NUMBER: 210121.439

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 206-622-4900

; TELEFAX: 206-682-6031

; INFORMATION FOR SEQ ID NO: 47:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3947 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; US-08-975-762-47

Query Match

Best Local Similarity 5.3%; Score 32.2; DB 4; Length 3947;

Matches 103; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 365 ttcttagtgcctgaggaagcaatgacgtgcaatccctcacattggttcacccag 424

Db 3375 TTTTGAATTTGATATATATGTTGTGTAGATTCCAGAGCTTTCTGTTTG 3434

QY 425 atcaaaactacatttggtgatlctcacaagaatgtagtgtaagggctgtactgtcc 484

Db 3435 ATGTTCTGCTAGTTGGTGGCTACCTGTGTAATCTATAGCTTTTATTTACACTTAAC 3494

QY 485 tgttcttggaatttgatcattgacgttcacgttcacattatgaaattcattctc 544

Db 3495 TAAATTTGAGAGTTGTGACGCAATTAATTTCTTAATTAATTAATTTGACCTTTTAAACCTCT 3554

QY 545 tgccttctcaatttgagggtgcacacagaggaattgcat 585

Db 3555 ACATATTCATTTCTTTTGTGAGGCTCTGATGACATGAT 3595

[illegible]

PS	Claim 13; SEQ ID 7155; 71pp + CD-ROM; English.	
XX		
CC	The present sequence is a polypeptide encoded by one of a large number	
CC	of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs	
CC	were prepared from total human RNAs or polyA+ RNAs derived from 30	
CC	different tissues. EST sequences usually correspond mainly to the 3'	
CC	untranslated region (UTR) of the mRNA because they are often obtained	
CC	from oligo-dT primed cDNA libraries. Such ESTs are not well suited for	
CC	isolating cDNA sequences derived from the 5' ends of mRNAs and even in	
CC	those cases where longer cDNA sequences have been obtained, the full 5'	
CC	UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'	
CC	ends and can therefore be used to obtain full length cDNAs and genomic	
CC	DNA. 5' ESTs are also used in diagnostic, forensic, gene therapy and	
CC	chromosome mapping procedures. They are used to obtain upstream	
CC	regulatory sequences and to design expression and secretion vectors.	
XX		
SO	Sequence 77 AA;	
XX		
Query Match	38.1%; Score 294; DB 21; Length 77;	
Best Local Similarity	83.6%; Pred. No. 2,3e-28;	
Matches 56; Conservative	3; Mismatches 8; Indels 0; Gaps	
OY	33 FSTOSPLOKLARKKKIIIGTIOILEGIMTFSGVFLFTLKPYRPFPPIFLSGVPFGS 92	
Db	11 fqlkapckkifarkmkligtltqxtlqimxfsigvflrlxlkxpprlpfllslxypfwg 70	
OY	93 VLFINSG 99	
Db	71 vlfinsg 77	
RESULT 2		
ID	AAVS3632 standard; Protein; 167 AA.	
XX	AAVS3632;	
XX	22-FEB-2000 (first entry)	
DE	A bone marrow secreted protein designated BMS208.	
XX		
KW	Bone marrow secreted protein; bone marrow stromal cell; cytokine;	
KW	cell proliferation; cell differentiation; hematopoiesis; anaemia;	
KW	myeloid cell deficiency; lymphoid cell deficiency; myeloid cell;	
KW	erythroid progenitor cell; colony stimulating factor; granulocyte;	
KW	monocyte; macrophage; myelo-suppression; megakaryocyte; platelet;	
KW	platelet disorder; thrombocytopenia; hematopoietic stem cell;	
KW	stem cell disorder; aplastic anaemia; bone differentiation;	
KW	paroxysmal nocturnal hemoglobinuria; bone growth; cartilage; tendon;	
KW	ligament; nerve; wound healing; tissue repair; burn; incision; ulcer;	
KW	bone fracture; cartilage damage; artificial joint.	
XX		
OS	Homo sapiens.	
XX		
XX	Key	Location/Qualifiers
FT	Peptide	1..62
FT	/note= "signal peptide"	
XX	WO9933979-A2.	
XX		
XX	08-JUL-1999.	
XX		
XX	18-DEC-1998;	98WO-US27008.
XX		
XX	30-DEC-1997;	97US-0068958.
XX	24-SEP-1998;	98US-0101603.
XX	30-SEP-1998;	98US-0102540.
XX		
XX	(CHIR) CHIRON CORP.	
XX		
XX	Lin H, Cao L;	
XX		

DR	WP1: 2000-038344/03.
DR	N-PSDB: AA236238.
XX	
PT	New isolated human polynucleotide and secreted proteins can induce
PT	production of other cytokines in certain cell populations -
XX	
PS	Claim 2; Page 94; 120pp: English.
XX	
CC	AAV53622-43 represent bone marrow secreted proteins of human bone marrow
CC	stromal cells. The proteins can exhibit cytokine, cell proliferation, or
CC	cell differentiation activity (either inducing or inhibiting). They can
CC	be used to support colony forming cells or factor-dependent cell lines,
CC	to regulate hematopoiesis, and to treat myeloid or lymphoid cell
CC	deficiencies. In addition, they may be used to support the growth and
CC	proliferation of erythroid progenitor cells, and to treat various
CC	anemias. They can have colony stimulating factor (CSF) activity and can
CC	be used to support the growth and proliferation of myeloid cells such as
CC	granulocytes, monocytes or macrophages, to prevent or treat
CC	myelo-suppression, to support the growth and proliferation of
CC	megakaryocytes and platelets, thereby allowing prevention or treatment
CC	of platelet disorders such as thrombocytopenia, to support the growth
CC	and proliferation of hematopoietic stem cells, either in place of or in
CC	conjunction with platelet transfusions, to treat stem cell disorders,
CC	such as aplastic anemia and pancytopenia, to treat hemoglobinuria, or to
CC	repopulate the stem cell compartment after irradiation or chemotherapy.
CC	They can be used for growth or differentiation of bone, cartilage,
CC	tendon, ligament, or nerve tissue, as well as for wound healing and
CC	tissue repair and replacement, and in the treatment of burns, incisions
CC	and ulcers, to induce cartilage and/or bone growth in circumstances
CC	where bone is not normally formed and thus have an application in healing
CC	bone fractures and cartilage damage or defects, prophylactic use in
CC	fracture reduction and also in the improved fixation of artificial
CC	joints.
XX	
SQ	Sequence 167 AA;
XX	
QY	Query Match 20.9%; Score 161.5; DB 21; Length 167;
	Best Local Similarity 35.6%; Pred. No. 8.4e-12;
	Matches 36; Conservative 20; Mismatches 44; Indels 1; Gaps 1.
Db	
39	LOKLAR-KMKLTGIIIOILFMGTFSGVIFLFLTLKPYRRPRFFLSGPPMGVLFIN 97
	: : : : : : : : : : : : : : : : : : : : :
35	Iqeklkgepkylgvyvqiltalmslsmjltumcmasntypnlsylygtlwwsvml 94
QY	98 SGAPLIAVRKRTTERILITLMTFSIIEFLFSLPFSIIGC 138
	: : : : : : : : : : : : : : : : : :
95	sgsisaagirtkylgldgmwllsvlefcavalsatgic 135
Db	
RESULT 3	
AA06503	
ID	AA06503 standard; Protein: 214 AA.
XX	
AC	AA06503;
XX	
DT	06-FEB-1997 (first entry)
XX	
DE	Htm4 protein.
XX	
XX	Human Htm4 protein; Fc(epsilon)RI receptor; FR, atopic disease;
KW	allergy; asthma; atopic dermatitis; allergic rhinitis; hereditary.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Domain
FT	Location/Qualifiers
FT	50..69 "Transmembrane domain"
FT	/note= "Transmembrane domain"
FT	89..108
FT	/note= "Transmembrane domain"
FT	121..141
FT	Domain
FT	/note= "Transmembrane domain"
FT	173..192
FT	Domain

XX 25-MAY-2000.
 PD 04-NOV-1999; 99MO-US26234.
 XX
 PF 18-NOV-1998; 98US-0195292.
 XX
 PR (INCY-) INCYTE PHARM INC.
 PA Walker MG, Volkmut W, Klingler TM;
 XX WPI: 2000-387787/33.
 DR N-PSDB; AAA27130.
 XX
 PT New human inflammation-associated polypeptide useful for diagnosis,
 PT prevention and treatment of inflammatory diseases comprises product of
 PT gene coexpressed with e.g. CD16, L-selectin and IP-30
 XX
 Claim 4; Page 39-40; 43pp; English.

CC Eleven novel inflammation-associated genes have been identified in
 CC cDNA libraries from various tissues. The genes were selected
 CC according to their coexpression with the known inflammation genes,
 CC CD16, L-selectin, Src-like adapter protein, IP-30, superoxide
 CC homoenzyme subunits, alpha-1-antitrypsin, C1q-A, 5-lipoxygenase
 CC activating protein and SRC family tyrosine kinase. The novel
 CC polynucleotides may be used in hybridization assays to diagnose a
 CC disease or condition associated with altered expression of the
 CC inflammation genes. Antibodies against the genes may be useful in
 CC compositions for the diagnosis and treatment of such diseases
 CC associated with inflammation including rheumatoid arthritis,
 CC Crohn's disease, multiple sclerosis, AIDS, diabetes mellitus,
 CC asthma and allergy. Additionally the polynucleotides of the
 CC invention may be used for gene therapy. The present sequence is
 CC human inflammation associated protein #6, derived from Incyte
 CC Clone 2349263.
 XX
 XX
 SQ Sequence 225 AA;

Query Match 18.0%; Score 139; DB 21; Length 225;
 Best Local Similarity 26.5%; Pred. No. 6.7e-09;
 Matches 41; Conservative 31; Mismatches 73; Indels 10; Gaps 4;

OY 1 MDSSTASHPVFLVPEPTITASEYESTELSATFTSTOSPLQKLPARKMKILGTIQLIFGIM 60
 1 mtsqpyneilivlpnvl--nfsqaekpeptngqdsllkhhakvlgltqilgmm 58
 61 TFSFGVIFLTLKP-YRPPFIPL-SGYPFGSVLFINSAGALLAVKRTTETL----- 113
 59 vlsigllasasfsnftqstclinsaypffigffllsqsistateklrlklyhssl 118
 Db
 OY 114 -GILITLMPFSIIEFLSPFSILGCHSEDCDEQ 147
 119 vgsilsalselvgfllsvkqactlnpsajqceldk 153
 Db

RESULT 6
 ID AAY73495 standard; Protein; 242 AA.
 XX AAY73495;
 AC
 XX
 XX 29-FEB-2000 (first entry)
 DT
 XX Human secreted protein clone y33_1 protein sequence SEQ ID NO:212.
 DE
 XX Human; secreted protein; immunostimulatory; haemostatic; cytokine;
 KW proliferative; differentiative; chemotactic; chemokinetic; vaccine;
 KW thrombolytic; antiinflammatory; cyostatic; immunosuppressive;
 KW gene therapy.
 XX
 OS Homo sapiens.

XX MO958642-A2.
 PN
 XX 18-NOV-1999.
 PD
 XX 14-MAY-1999; 99MO-US10843.
 PF
 XX 14-MAY-1998; 98US-0085472.
 PR 17-AUG-1998; 98US-0096824.
 PR 11-SEP-1998; 98US-0099843.
 PR 11-SEP-1998; 98US-0099950.
 PR 15-SEP-1998; 98US-0100424.
 PR 29-SEP-1998; 98US-0102329.
 PR 09-OCT-1998; 98US-0103615.
 PR 11-DEC-1998; 98US-0111799.
 PR 14-DEC-1998; 98US-0112159.
 PR 31-DEC-1998; 98US-0114415.
 PR 10-FEB-1999; 99US-0248059.
 PR 06-APR-1999; 99US-0287150.
 PR 13-MAY-1999; 99US-0311021.
 XX

PA (GEM) GENETICS INST INC.
 XX Wong GG, Clark HF, Fechtel K, Agostino MJ;
 PI WPI: 2000-053095/04.
 DR N-PSDB; AA252580.
 XX
 PT Novel polynucleotides and proteins having biological activities which
 PT make them suitable for treating, preventing or ameliorating medical
 PT conditions in humans or animals
 XX
 PS Claim 221; Page 718; 730pp; English.

CC The present invention describes human secreted proteins encoded by
 CC polynucleotides obtained from adult testes, foetal brain, adult brain,
 CC brain (foetal and adult), foetal kidney, adult spleen, and adult thymus
 CC cDNA libraries. The polynucleotides and proteins are predicted to have
 CC biological activities which would make them suitable for treating,
 CC preventing or ameliorating medical conditions in humans and animals.
 CC Suggested activities include nutritional activity, cytokine and cell
 CC proliferation/differentiation activity, immune stimulating (e.g. as
 CC vaccines) or suppressing activity, haematopoiesis regulating activity,
 CC tissue growth activity, activin/inhibin activity, chemotactic/
 CC chemokinetic activity, haemostatic and thrombolytic activity, receptor/
 CC ligand activity, anti-inflammatory activity, cadherin/tumour invasion
 CC suppressor activity, and tumour inhibition activity. The polynucleotides
 CC are also stated to be useful for gene therapy. Therapeutic compositions
 CC are also presently valuable for veterinary applications. AA252475 to
 CC AA252581 encode human secreted proteins, and AAY73390 to AAY73500
 CC represent human secreted proteins, given in the present invention.
 XX

SQ Sequence 242 AA;

Query Match 18.0%; Score 139; DB 21; Length 242;
 Best Local Similarity 26.5%; Pred. No. 7.4e-09;
 Matches 41; Conservative 31; Mismatches 73; Indels 10; Gaps 4;

OY 1 MDSSTASHPVFLVPEPTITASEYESTELSATFTSTOSPLQKLPARKMKILGTIQLIFGIM 60
 1 mtsqpyneilivlpnvl--nfsqaekpeptngqdsllkhhakvlgltqilgmm 58
 61 TFSFGVIFLTLKP-YRPPFIPL-SGYPFGSVLFINSAGALLAVKRTTETL----- 113
 59 vlsigllasasfsnftqstclinsaypffigffllsqsistateklrlklyhssl 118
 Db
 OY 114 -GILITLMPFSIIEFLSPFSILGCHSEDCDEQ 147
 119 vgsilsalselvgfllsvkqactlnpsajqceldk 153
 Db

RESULT 7

AAV15225
ID AAV15225 standard; protein; 248 AA.
XX
AC AAV15225;
XX
DT 26-OCT-1999 (first entry)
XX
DE Human receptor protein (HURP) 4 amino acid sequence.
XX
KW receptor; cancer; autoimmune disorder; inflammation;
KW antagonist; cell surface protein; cell signalling;
KW antibody; human receptor protein; HURP; reproductive disorder;
KW developmental disorder; gastrointestinal disorder.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 8 /note= "Potential N-glycosylation site"
FT Misc-difference 20 /note= "Potential N-glycosylation site"
FT Misc-difference 74 /note= "Potential N-glycosylation site"
FT Misc-difference 110 /note= "Potential N-glycosylation site"
FT Misc-difference 22 /note= "Potential CAMP-/cGMP-dependent protein-kinase phosphorylation site"
FT Misc-difference 193 /note= "Potential casein kinase II-phosphorylation site"
FT Misc-difference 193 /note= "Potential casein kinase II-phosphorylation site"
FT Misc-difference 36 /note= "Potential protein kinase C-phosphorylation site"
FT Misc-difference 105 /note= "Potential protein kinase C-phosphorylation site"
FT Misc-difference 136 /note= "Potential protein kinase C-phosphorylation site"
FT Misc-difference 177 /note= "Potential protein kinase C-phosphorylation site"
FT Misc-difference 233 /note= "Potential protein kinase C-phosphorylation site"
FT Misc-difference 242 /note= "Potential tyrosine kinase phosphorylation site"
PN WO941375-A2.
XX
PD 19-AUG-1999.
XX
PF 05-FEB-1999; 99WO-US02572.
XX
PR 12-FEB-1998; 98US-0022939.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Au-Young J, Bandman O, Baughn M, Corley NC, Guegler KJ;
PI Hillman JL, Lal P, Shah P, Tang YT, Yue H;
DR WPI; 1999-494536/41.
DR N-PSDB; AA206369.
XX
PT New human receptor proteins, used e.g. to treat, prevent and
PT diagnose gastrointestinal and developmental disorders - and related
PT nucleic acids, vectors, transformed cells, antibodies, agonists and
PT antagonists
XX
PS Claim 1; Page 81-82; 94pp; English.

XX
CC The Human receptor protein 4 (HURP-4) has 22% homology with the
CC rat 19E receptor and 19% homology with human CD20 protein.
CC HURP-4 is expressed in cancerous, inflamed, hematopoietic/immune
CC and gastrointestinal tissue. HURP-4 therefore appears to have a role in
CC some forms of cancer, autoimmune/inflammatory disorders, and
CC gastrointestinal disorders.
CC This gives rise to the possibility of using an antagonist or an antibody
CC of HURP to treat or prevent cancer or autoimmune/inflammatory disorders.
XX
SQ Sequence 248 AA;
XX
Query Match 18.0%; Score 139; DB 20; Length 248;
Best Local Similarity 26.5%; Pred. NO. 7.6e-09;
Matches 41; Conservative 31; Mismatches 73; Indels 10; Gaps 4;
OY 1 MDSSTANSPVFLVPPETITASEVESTELSAVTFSTQSPQKLFARKMKITGTOILFGIM 60
DB 1 mtsqpyneclivpsnvl--nfsqaekpeptngqslkhlhaelkvgtlqlcgmm 58
OY 61 TFSRGVIFLFTLKP-YRPPFIFL-SGYPPWGSVLFINSGAFILAVKRTTEFL----- 113
DB 59 vlsigllasasfapnftqvtstllnsayftgpfiffisgslatekrtikilvhs1 118
OY 114 -GILITLMTFSLIFLISLPSILGCHSEDDCEQ 147
DB 119 vgsllsalsalvgfllsvkqatlnpaslqceldk 153
XX
RESULT 8
AAV1531
ID AAV1531 standard; protein; 248 AA.
XX
AC AAV1531;
XX
DT 29-JUN-2000 (first entry)
XX
DE Human secreted protein sequence encoded by gene 81 SEQ ID NO:204.
XX
KW Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
KW antiHIV; antiinflammatory; neotropic; neuroprotective; antiasthma;
KW osteopathic; antiarthritic; antibacterial; antidiabetic; antisthma;
KW antiparasitic; cardiac; gene therapy; cancer; neurological disorder;
KW immune disease; inflammation; blood disorder; tumour; chromosome 1.
XX
OS Homo sapiens.
XX
PN WO200006698-A1.
XX
PD 10-FEB-2000.
XX
PF 29-JUL-1999; 99WO-US17130.
XX
PR 30-JUL-1998; 98US-0094657.
PR 05-AUG-1998; 98US-0095486.
PR 06-AUG-1998; 98US-0095454.
PR 06-AUG-1998; 98US-0095455.
PR 12-AUG-1998; 98US-0096319.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Komatsoulis GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;
PI Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;
PI Soppet DR, Endress GA, Ebner R, Olsen HS, Mucenski M;
DR WPI; 2000-195282/17.
DR N-PSDB; AAA26426.
XX
PT New isolated human genes and the secreted polypeptides they encode,
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders -
XX

PS Claim 11; Page 513-514; 634pp; English.

XX The polynucleotide sequences given in AAA26346 to AAA26458 encode the
CC human secreted proteins given in AAY91451 to AAY91691. The human secreted
CC proteins can have activities based on the tissues and cells they are
CC expressed in. Examples of the activities are: cytostatic;
CC immunosuppressive; antiHIV; antiinflammatory; nootropic; neuroprotective;
CC antiarthritis; osteoporosis; antibacterial; antidiabetic;
CC antiasthma; antipsoriatic; and cardiant. The polynucleotides and their
CC corresponding secreted proteins are useful for preventing, treating or
CC ameliorating medical conditions, e.g. by protein or gene therapy. Also
CC pathological conditions can be diagnosed by determining the amount of the
CC proteins in a sample or by determining the presence of mutations in the
CC polynucleotides. Specific uses are described for each of the
CC polynucleotides, based on which tissues they are most highly expressed
CC in, and include developing products for the diagnosis or treatment of
CC cancer, tumours, neurodegenerative disorders, developmental abnormalities
CC and foetal deficiencies, blood disorders, diseases of the immune system,
CC autoimmune diseases, hepatic and renal disorders, inflammation,
CC allergies, Alzheimer's and behavioural disorders, schizophrenia,
CC osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
CC transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
CC cardiovascular disorders, reproductive disorders, gastrointestinal
CC disorders, respiratory disorders and metabolic disorders. The proteins
CC or polynucleotides can also be used as food additives or preservatives.
CC The proteins are also useful for identifying their binding partners.
CC AAA26337 to AAA26345 and AAY91450 are sequences used in the
CC exemplification of the present invention.

SO Sequence 248 AA;

Query Match 18.0%; Score 139; DB 21; Length 248;
Best Local Similarity 26.5%; Pred. No. 7.6e-09;
Matches 41; Conservative 31; Mismatches 73; Indels 10; Gaps 4;

QY 1 MDSSTASPVFLVPPPEITASEYESTELSTSTQSPLOKLFARKKKIIGTIIQILFGIM 60
DB 1 mtsqpyneclivlpnsnvi--nfsqaekpeptngqdsllkhhaelkvigtqilqgmm 58
QY 61 TFSFGVIFLTLKP-YPRPEIFL-SGYPPWGSVLFINSGAFLIAVKKRTETL----- 113
DB 59 vlsigillaasfsfntqyvtstlinsaypfifgffifligsistatekrltklilvns1 118
QY 114 -GILITMTFSIIEFLISLPSILGCHSEDCDCQ 147
DB 119 vgsilsalsalvgfllsvkqatlnpsalqceldk 153

RESULT 9

AAB70489 standard; Protein; 248 AA.

XX AAB70489;

DT 04-MAY-2001 (first entry)

DE Human hHAIRBs-iso protein sequence SEQ ID NO:7.

XX Human; hHAIRBs-iso; HAIRBs; HAIRBs isomer; detection;

KW high affinity immunoglobulin epsilon receptor beta subunit.

OS Homo sapiens.

XX CNI269410-A.

PD 11-OCT-2000.

PE 17-MAR-2000; 2000CN-0114959.

PR 17-MAR-2000; 2000CN-0114959.

PA (SREN-) SOUTHERN RES CENT NAT HUMAN GENE GROUP.

XX Xiao H, Liu F, Song H;

DR WPI; 2001-050545/07.

XX N-PSDB; AAF63724.

PT New human immunoglobulin receptor subunit protein and its nucleic acid

PS Claim 4; Page 17; 22pp; Chinese.

XX The present invention describes a human high affinity immunoglobulin
CC epsilon receptor beta subunit isomer, designated hHAIRBs-iso.
CC hHAIRBs-iso is isolated from in human pheochromocytoma. The present
CC invention also describes methods for the preparation and detection of
CC hHAIRBs-iso protein and nucleotide sequences. The present sequence
CC represents the human hHAIRBs-iso protein, as given in the present
CC invention.

SO Sequence 248 AA;

Query Match 18.0%; Score 139; DB 22; Length 248;
Best Local Similarity 26.5%; Pred. No. 7.6e-09;
Matches 41; Conservative 31; Mismatches 73; Indels 10; Gaps 4;

QY 1 MDSSTASPVFLVPPPEITASEYESTELSTSTQSPLOKLFARKKKIIGTIIQILFGIM 60
DB 1 mtsqpyneclivlpnsnvi--nfsqaekpeptngqdsllkhhaelkvigtqilqgmm 58
QY 61 TFSFGVIFLTLKP-YPRPEIFL-SGYPPWGSVLFINSGAFLIAVKKRTETL----- 113
DB 59 vlsigillaasfsfntqyvtstlinsaypfifgffifligsistatekrltklilvns1 118
QY 114 -GILITMTFSIIEFLISLPSILGCHSEDCDCQ 147
DB 119 vgsilsalsalvgfllsvkqatlnpsalqceldk 153

RESULT 10

AAV48505 standard; Protein; 250 AA.

XX AAV48505;

DT 08-DEC-1999 (first entry)

DE Human breast tumour-associated protein 50.

XX Expressed sequence tag; EST; human; breast; cancer; cytostatic;

KW medicaments; gene therapy; treatment; fat metabolism.

OS Homo sapiens.

PN DE19813835-A1.

PD 23-SEP-1999.

PE 20-MAR-1998; 98DE-1013835.

PR 20-MAR-1998; 98DE-1013835.

XX (META-) METAGEN GES GENOMFORSCHUNG MBH.

PA Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;

DR WPI; 1999-528979/45.

XX N-PSDB; AAZ33586.

PT Human nucleic acid sequences and protein products from normal breast

PS Claim 28; 180; 206pp; German.

[illegible]

XX	Key	Location/Qualifiers
XX	Modified-site	125
XX	Modified-site	/note= "potential protein kinase C phosphorylation site"
XX	Modified-site	137
XX	Modified-site	/note= "potential N-glycosylation site"
XX	Modified-site	192
XX	Modified-site	/note= "potential casein kinase II phosphorylation site"
XX	US5871930-A.	
XX	16-FEB-1999.	
XX	21-AUG-1997;	97US-0916902.
XX	21-AUG-1997;	97US-0916902.
XX	(INCY-) INCYTE PHARM INC.	
XX	Bandman O, Corley NC, Lal P;	
XX	WPI; 1999-166628/14.	
XX	N-PSDB; AAX15104.	
XX	New polynucleotide coding for high affinity immunoglobulin E receptor-like protein - and its complement, useful for diagnosis, prevention and antisense therapy of inflammatory responses	
XX	Claim 1; Fig 1A-D; 30pp; English.	
XX	The present sequence represents a high affinity immunoglobulin E receptor-like protein (IGERB). The present sequence was first identified in Incyte clone 927955 for the brain cDNA library BRAIN0704. Polynucleotides complementary to the IGERB cDNA can be used as probes to IGERB gene expression. IGERB proteins and nucleotides may be used for the diagnosis of inflammatory responses associated with expression of IGERB, e.g AIDS, Addison's disease, atherosclerosis, bronchitis, ulcerative colitis, diabetes mellitus, emphysema, gout, Graves's disease, osteoporosis, rheumatoid arthritis, Sjogren's syndrome and complications of cancer. The sequences may also be useful in assays that detect activation or induction of various cancers.	
XX	Sequence	239 AA;
XX	Query Match	17 4%; Score 134; DB 20; Length 239;
XX	Best Local Similarity	31.4%; Pred. No. 2.9e-08;
XX	Matches	38; Conservative 20; Mismatches 47; Indels 16; Gaps 3
XX	39 LOKLPR-KMKILGTQILGIMTFSEGVIFFTLLKPYRPFIFLISGYPFGSVLEFIN 97	
XX	DB 54 lqekfigspkrlgyvvqlltalsmsjgmllmcmsantysgnplsvlytllgswmfil 113	
XX	QY 98 SCAFLIAVKRKTE-----TLGILTF-----LMFSLIEFLISLPFSLGCHSED 142	
XX	DB 114 sgslsaaqirtkglvrgslgmaltssvlaasgillnftslafysfhbpcnygnsmn 173	
XX	QY 143 C 143	
XX	DB 174 C 174	
XX	RESULT 15	
XX	AAAY50174	
XX	AAAY50174 standard; Protein: 239 AA.	
XX	AAAY50174;	
XX	31-JAN-2000 (first entry)	

DE			Human hIgh affinity Ige receptor-like protein (IGERB).
XX	IGERB; high affinity; immunoglobulin E; IGE; receptor; allergy;		
KM	Inflammation; B lymphocyte; homology; high affinity IGE receptor; FCRI;		
KW	mast cell; basophil; histamine; protease; cytoplasmic granule; synthesis;		
KV	effector; prostaglandin; leukotriene; cytokine; antigen; antibody;		
KM	identification; agonist; antagonists; expression; activity; diagnosis;		
KK	therapy; inflammatory disorder; multiple sclerosis; osteoarthritis;		
XX	asthma; cancer; side effect; complication.		
OS	Homo sapiens.		
FH	Key	Location/Qualifiers	
FT	Modified-site	125	/note= "Phosphorylated by protein kinase C"
FT	Modified-site	137	/note= "Glycosylated"
FT	Modified-site	192	/note= "Phosphorylated by casein kinase II"
XX			
PN	US5977072-A.		
PD	02-NOV-1999.		
XX			
PF	15-DEC-1998;	98US-0213389.	
PR	21-AUG-1997;	97US-0916902.	
PA	(INCY-) INCYTE PHARM INC.		
PI	Bandman O, Corley NC, Lal P;		
DR	WPI; 2000-012123/01.		
N-PSDB:	AZ32842.		
PT	High affinity immunoglobulin E receptor-like protein useful for stimulating allergic and immune responses -		
Claim 1;	Fig 1; 29pp; English.		
This sequence represents human high affinity immunoglobulin E (IGE)			
receptor-like protein (IGERB). Nucleic acids encoding IGERB were			
initially identified in a brain cDNA library, this sequence being a			
consensus. An allergic response is initiated by release of IGE			
from B lymphocytes. The IGE molecules then bind to the high affinity			
IGE receptor (FcR) present on mast cells and basophils, which triggers			
the release of histamine and proteases from cytoplasmic granules and			
leads to the synthesis of effectors of the allergic and inflammatory			
response, such as prostaglandins, leukotrienes and cytokines. As IGERB			
binds IGE, it may be administered to stimulate allergic and immune			
responses in patients in whom IGERB is under expressed or inactive and			
to supplement the patient's own production of the protein. IGERB may also			
be used as an antigen for the production of antibodies and to identify			
candidate agonists and antagonists of IGERB expression and activity.			
The antibodies may also be used in diagnosis. Antibodies and antagonists			
may be administered to downregulate IGERB activity and reduce the			
potency of inflammatory and allergic responses. They may be used in this			
way to treat inflammatory disorders such as multiple sclerosis,			
osteoarthritis, asthma and some complications of cancer. Conversely, the			
agonists may be used to enhance immune responses.			
Sequence 239 AA:			
Query Match	17.4%;	Score 134;	DB 21; Length 239;
Best Local Similarity	31.4%;	Pred. No. 2.9e+08;	
Matches 38;	Conservative 20;	Mismatches 47;	Indels 16; Gaps 3;
39 LQLEPARKMLITGLTQLFGIMFSGYIVLFLTLKPYPFRFFILSGYPMGSVLFIN 97			
: : : : : : : : : : : : : : : : :			
Db 54 IQEKLKGEPKVLYGVQTALTSMSGTMMCMASNYCSPISLVYGYLTGWSEVMFI 113			
98 SGAFIAAVKRRTTE-----TLGITL-----LMTFSLIEELFSIPSTLGCHSD 142			

Thu Oct 11 10:24:08 2001

us-09-735-712-8.rag

Page 10

Db 114 sgs1s1aaglrctckgylvrsglgnmtssvlaasgllntfslafysfhpypcnyygnsnn 173
QY 143 C 143
Db 174 c 174

Search completed: October 7, 2001, 04:39:26
Job time: 1911 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2001, 04:40:00 ; Search time 24.66 Seconds

(without alignments)
124.410 Million cell updates/sec

Title: US-09-735-712-8

Sequence: 1 MDSSTAHSPVFLVPPETIA.....SLPFSILGCHSEDCDCGCC 149

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	150.5	19.5	214	1	US-08-318-492-4
2	150.5	19.5	214	1	US-08-707-340-4
3	150.5	19.5	214	1	US-08-916-902A-3
4	150.5	19.5	214	2	US-08-994-578-4
5	150.5	19.5	214	2	US-09-213-389-3
6	134	17.4	239	2	US-08-916-902A-1
7	134	17.4	239	2	US-09-213-389-1
8	123.5	16.0	243	1	US-07-869-933-29
9	123.5	16.0	243	1	US-07-869-933-33
10	123.5	16.0	243	1	US-08-201-879A-4
11	123.5	16.0	243	2	US-08-916-902A-4
12	123.5	16.0	243	2	US-09-213-389-4
13	123.5	16.0	243	4	US-09-103-663-29
14	123.5	16.0	243	4	US-09-103-663-33
15	123.5	16.0	246	1	US-07-869-933-23
16	123.5	16.0	246	4	US-09-103-663-23
17	117.5	15.2	235	1	US-07-869-933-34
18	117.5	15.2	235	1	US-08-201-879A-5
19	117.5	15.2	235	1	US-09-103-663-34
20	99.5	12.9	244	1	US-07-869-933-32
21	99.5	12.9	244	4	US-08-201-879A-3
22	99.5	12.9	244	4	US-09-103-663-32
23	75	9.7	325	5	PCR-US93-08528-30
24	72.5	9.4	360	4	US-08-875-573-20
25	72.5	9.4	360	4	US-09-232-878-2
26	72.5	9.4	360	4	US-09-232-878-2
27	68	8.8	483	1	US-08-194-338-7

28	65.5	8.5	591	1	US-08-484-840-2	Sequence 2, App11
29	65.5	8.5	591	1	US-08-483-094-2	Sequence 2, App11
30	64	8.3	215	4	US-09-087-232A-17	Sequence 17, App1
31	64	8.3	239	2	US-08-933-750C-22	Sequence 22, App1
32	64	8.3	239	4	US-09-234-613-22	Sequence 22, App1
33	64	8.3	329	2	US-09-087-232A-13	Sequence 13, App1
34	63.5	8.2	359	2	US-08-749-289-1	Sequence 2, App11
35	63	8.2	549	2	US-08-791-887-2	Sequence 2, App11
36	63	8.2	603	4	US-09-097-889-23	Sequence 23, App1
37	62	8.0	376	2	US-08-875-972-2	Sequence 2, App11
38	62	8.0	1353	3	US-08-894-173-2	Sequence 2, App11
39	62	8.0	1353	4	US-09-398-193-2	Sequence 9, App11
40	61.5	8.0	365	2	US-08-467-559B-9	Sequence 9, App11
41	61	7.9	235	2	US-08-580-545B-10	Sequence 10, App1
42	61	7.9	235	4	US-09-262-653A-10	Sequence 10, App1
43	61	7.9	1024	4	US-09-091-117-5	Sequence 5, App11
44	60.5	7.8	855	6	5196526-1	Patent No. 5196526
45	60.5	7.8	1471	1	US-08-683-839B-3	Sequence 3, App11

ALIGNMENTS

```
RESULT 1
US-08-318-492-4
; Sequence 4, Application US/08318492
; Patent No. 5552312
;
GENERAL INFORMATION:
; APPLICANT: Lim, Bing
; APPLICANT: Adra, Chaker N.
; APPLICANT: Lelias, Jean-Michel
; TITLE OF INVENTION: RECOMBINANT HUM4 GENE, PROTEIN AND
; TITLE OF INVENTION: ASSAYS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millitia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02173
;
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
;
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,492
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: BIH94-03
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 214 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-318-492-4
;
Query Match 19.5%; Score 150.5; DB 1; Length 214;
Best local Similarity 31.1%; Pred. No. 2,1e-10;
Matches 51; Conservative 22; Mismatches 64; Indels 27; Gaps 8;
```

QY 1 MDSSTAHSPVFLVPPETIASSESTELSATTFSTQSPLOKLFARKMKITGLTQILFGI 59
Db 11 LGSASAHGPGSGSETGPELINTSYH-----PINGSPPDYOK--AKIQLVLAIGIILNAA 60

[illegible]

Query Match 19.5%; Score 150.5; DB 2; Length 214;
Best Local Similarity 31.1%; Pred. No. 2.le-10;
Matches 51; Conservative 22; Mismatches 64; Indels 27; Gaps 8;

Dy 1 MDSSYAH-SPEVLVPPETASSEYSTELSATFTFSQSLPKLFAKKKILGCTIOLFGI 59
: : : : : | : : : : : | : : : : :
Db 1 LGSABHGTFCGSEICPELTNLSVH-----PINGSDDYCK--AKLVLCATIDILNNA 60
60 MTFSFGLVFLEPTLLKPYP---PRPFIFLGGVPFWGSLVINSGAFLIAVKRKTET--- 112
| : : : : : | : : : : : | : : : : : | : : : : :
Db 61 MILAGV-FLSLQPYHFQKHFFFTTYTGPIGVANGAFPCSSGTLNVAGIKPRTNIQ 119
| : : : : : | : : : : : | : : : : : | : : : : :
Cy 113 --LGILLMTFSIT-ELFISLPF----SLIGCHSEDCDEOC 148
| : : : : : | : : : : : | : : : : : | : : : : :
Db 120 NSFGMINASATIALVTGTAFLSLINIAVNIOQLSRCHSSESSESDLC 163

RESULT 6
US-08-916-902A-1
Sequence 1, Application US/08916902A
Patent No. 5871930
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Incyte Pharmaceuticals, Inc.

```

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/213,389
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/916,902
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PP-0371 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 239 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINOT04
CLONE: 927955
US-09-213-389-1

Query Match      17.4%; Score 134; DB 2; Length 239;
Best Local Similarity 31.4%; Pred. No. 2,38-08;
Matches 38; Conservative % 20; Mismatches 47; Indels 16; Gaps 3;

QY   39 LOKLFAR-KMLILGIQLIFGIMTFSGVITFLTKYPFRPFLLISGYFWGSYLFIN 97
    || : : : || : ||| : : | : : | : ||||| : ||
Db   54 LQEKLKPEPYLVGVQILTALMSLSKGITMCMAASNTGSPISVYGIVGWSMVFII 113
QY   98 SGAPLIAVKRRKYTE-----TLGILIT-----LMTFSIIEFLISLPSSIIAGHSED 142
    || : : : ||| : : || : ||| : | : ||| : | :
Db   114 SGLSIAGINTPTKKLVKSLGMNITSSVLAAAGILMTFSLAFPSFHHPYCNYGNNSNN 173
QY   QY   143 C 143
    Db     | - F
        174 C 174

RESULT      8
US-07-869-933-29
Sequence 29, Application US/07869933
Patent No. 5770396
GENERAL INFORMATION:
APPLICANT: KINET, Jean-Pierre
TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF
TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR
TITLE OF INVENTION: IMMUNOGLOBULIN
NUMBER OF SEQUENCES: 34.
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/869,933
FILING DATE: 19920416
```



```

J. Immunol. 144:4388-4394, 1988
J>Title: Cloning of a complementary DNA encoding a new mouse B lymphocyte differentiation antigen
A:Reference number: A30558; MUID:89067519
A:Accession: A30558
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-291 <TED>
C:Genetics:
A:Map position: 19
C:Superfamily: B-cell surface antigen CD20
C:Keywords: membrane protein; phosphoprotein; surface antigen

Query Match          16.9%; Score 172; DB 2; Length 291;
Best Local Similarity 29.4%; Pred. No. 7e-09;
Matches 53; Conservative 29; Mismatches 64; Indels 34; Gaps 7;

OY      32 TESTSQPLKLFERKKMLIGTIIQLIFGIMTFSEFGVIFLFTLLKPYPREFIFLS-GYPEW 90
       | | | | | | | | | | | | | | | | : : : : : | | | | |
27 TSSLVGPLOSFFMRRESALANQVMNGILFHTTLGL-----LMIPGVFAPICLSWYPLM 82
       | : : | | : : : : : : : : : : | | | | |
DB      91 GSVLFINSGAFLLAVKRKRTETLTILSR-IMNLSALRALAGIIL-----LTFEGFIID 142
       | : : | | : : : : : : : : : : | | | | |
DB      83 GGIMYITISGSLAAAEKTSRKSLVKAKVIMSLSLFLPAISGIILSIMDILNNITLSHFELK 142
       | : : | | : : : : : : : : : : | | | | |
OY      143 QN-----YIC---GYSHONS-----QCKATVLEFLGITLMTSIIELFIS 181
       | | | | | | | | | | | | | | | | : : : : :
DB      143 MRRLIELIQSKPYVDIYDCERSNSEKNPSSTOYCNSIOSVFGLISAWLISAFOKLVT 202
       | | | | | | | | | | | | | | | | : : : : :

RESULT 3
B-cell surface antigen CD20 - human
N:Alternate names: B-lymphocyte antigen CD20; B1
C:Species: Homo sapiens (man)
C:Date: 08-Jun-1989 #sequence.revision 02-Aug-1996 #text_change 22-Jun-1999
C:Accession: A30586; JI0042; A27400; S00387
R:Redder, T.F.; Klejman, G.; Schlossman, S.F.; Saico, H.
J. Immunol. 142, 2560-2568, 1989
A>Title: Structure of the gene encoding the human B lymphocyte differentiation antigen C
A:Reference number: A30586; MUID:89176281
A:Accession: A30586
A:Molecule type: DNA
A:Residues: 1-297 <TED>
A>Note: The authors translated the codon ATG for residue 148 as His, TCA for residue 289 and GAC for residue 294 as Ser
Stamenkovic, I.; Seed, B.
Exp. Med. 167, 1975-1980, 1988
Title: Analysis of two cDNA clones encoding the B lymphocyte antigen CD20 (B1.Bp35), a
A:Reference number: JI0042; MUID:88258386
A:Accession: JI0042
A:Molecule type: mRNA
A:Residues: 1-297 <STA>
A:Cross-references: GB:X12530; NID:g29773; PIDN:CAA31046.1; PID:g29774
R:Redder, T.F.; Streuli, M.; Schlossman, S.F.; Saico, H.
Proc. Natl. Acad. Sci. U.S.A. 85, 208-212, 1988
A>Title: Isolation and structure of a cDNA encoding the B1 (CD20) cell-surface antigen C
A:Reference number: A27400; MUID:88124792
A:Accession: A27400
A:Molecule type: mRNA
A:Residues: 1-297 <TE2>
A:Cross-references: GB:M27394; GB:J03574; NID:g179307; PIDN:AAA35581.1; PID:g179308
Reinfeld, D.A.; Brown, J.P.; Valentine, M.A.; Clark, E.A.; Ledbetter, J.A.
EMBO J. 7, 711-717, 1988
A>Title: Molecular cloning of the human B cell CD20 receptor predicts a hydrophobic prot
A:Reference number: S00387; MUID:88283639
A:Accession: S00387
A:Molecule type: mRNA
A:Residues: 1-12, 'L', 'L', 'L', 'L' <EIN>
A:Cross-references: EMBL:X07203; NID:g29775; PIDN:CAA30179.1; PID:g29776
A:Experimental source: David cells
C:Comment: This protein appears not to be glycosylated. Isoforms of varying molecular we
C:Comment: This protein plays an important role in B cell activation.
```

```

C.Genetics:
A:Gene: GDB:CD20
A:Cross-references: GDB:119761; OMIM:112210
A:Map position: 11q12-11q13.1
C:Superfamily: B-cell surface antigen CD20
C:Keywords: B-cell; phosphoprotein; transmembrane protein
F:1-51/Domain: Intracellular #status predicted <CYT>
F:52-103/Domain: transmembrane #status predicted <TM12>
F:104-116/Domain: Intracellular #status predicted <CYT7>
F:117-141/Domain: transmembrane #status predicted <TM3>
F:142-185/Domain: extracellular #status predicted <EX7>
F:186-212/Domain: transmembrane #status predicted <TM4>
F:213-297/Domain: intracellular #status predicted <CYT3>
F:167-183/Distulfide bonds: #status predicted

Query Match          16.6%; Score 168; DB 1; Length 297;
Best Local Similarity 30.3%; Pred. No. 1.7e+08;
Matches 54; Conservative 27; Mismatches 63; Indels 34; Gaps 7;

QY 34 STQSPLOKIFARKMKILGTIQLFGIMTFPSFGVIFLTLKPYPRE-PFILSGYPMGS 92
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 36 SLVGPTQSFFMRSKTKLGAVQIMNGLFHIALGCL-----LMIPACIVADICVTWYPLMGC 91

QY 93 VLFINSGAFLIAVKRKTTETLLISRMNLASALPAINGIILL-----T 136
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 92 IMYIIIGSILATEKNSRKCIVKGKRMINMSLSLFAISGMILSTMDLINIKISHFLKMS 151

QY 137 FGFILDON-YICGY-----SHONS----CKAAVTVLEFLITLMTFSII-ELFIS 181
      ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 152 LMFIRAHPIYINITYCEPANSPEKNSPSTQYCYISQSLFGLISVMILFAFPQELVIA 209

RESULT 4
A1231
high-affinity Ige receptor beta chain - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 05-Nov-1999
C:Accession: A31231
R.Kinet, J.P.; Blank, U.; Ra, C.; White, K.; Metzger, H.; Kochan, J.
Proc. Natl. Acad. Sci. U.S.A. 85, 6483-6487, 1988
A>Title: Isolation and characterization of cDNAs coding for the beta-subunit of the h
A:Reference number: A31231; MUID:88320465
A:Accession: A31231
A:Molecule type: mRNA
A:Residues: 1-243 <KIN>
A:Cross-references: GB:M22923; GB:J03845; NID:g204116; PIDN:AAA41149.1; PID:g204117
C:Keywords: immunoglobulin receptor; transmembrane protein

Query Match          15.7%; Score 159; DB 2; Length 243;
Best Local Similarity 28.0%; Pred. No. 9.9e+08;
Matches 45; Conservative 29; Mismatches 71; Indels 16; Gaps 3;

QY 36 QSPLOKLARRKKKILGTIQLFGIMTFPSFGVIFLTLKP--YPRPFILSGYPMGSV 93
      | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db 47 QQTWSQFLKELEFGVQVLVGLCLCFGVYCSTLTQSDFEDEVLLVLRAGYPMGAV 106

QY 94 LFINSGAFLIAVKRKTTETLLISRMNLASRLRIAGIITLFTGFIIDQNYICYGNQN 153
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 107 LVVLSGFLISMSEKNTLYLVKSGISGANIVSSIAAGLAIALLNLNSNSAYM----- 159

QY 154 SOCKAVT-----VLFLGIITLMTFSIIIEFLISLPESIL 187
      : | | : | | : | | : | | : | | : | | : | | : | | : | | : | |
Db 160 NVCKDITEDDCGVYSFITTELVLMLFLTLLAFCSAVLLII 200

RESULT 5
B34342
IGE Fc receptor beta chain - mouse
C:Species: Mus musculus (house mouse)
C:Date: 15-Jun-1990 #sequence_revision 15-Jun-1990 #text_change 05-Nov-1999
C:Accession: B34342
```

R:Ra, C.; Jouvin, M.H.E.; Kinet, J.P.
 J. Biol. Chem. 264, 15323-15327, 1989
 A:Title: Complete structure of the mouse mast cell receptor for IgE (re-epsilon-RI) and
 A:Reference number: A34342; MUID:89359361
 A:Accession: B34342
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-235 <RAC>
 A:Cross-references: GB:J05019; NID:g193238; PIDN:AAA37601.1; PID:g309225
 C:Keywords: Immunoglobulin receptor; Transmembrane protein

Query Match 15.6%; Score 158.5; DB 2; Length 235;
 Best Local Similarity 30.5%; Pred. No. 1.1e-07;
 Matches 60; Conservative 29; Mismatches 81; Indels 27; Gaps 9;

QY 1 MDSSTAHPVFLVFP-PEITASEYESTELSATPTQSPLO-KLPARK-MKLTGTIQL 56
 1 MDREN-RSRADLAMPQESSAPDIELLEASPAKAPKQMTREFLKELEFGATQIL 59
 DB 1
 57 FGIMTFEFGYIFLTLKPR-----PRPFIFLSGYFPGSVLFINGATLAVKRTTE 111
 60 VGLICCFEFGIVSVL--YVSPDEEVLLLYKGYEPFGVGLFVLSGFLSIISERNNTL 116
 QY 112 TLILSRIMMLSLRAIAGIILTFEGFILDONVICGYSHONSCKAVT-----VLFL 164
 117 YLVNGSIGANIVSSIAAGTGMIL-----NLNNFAIMNN-CKNVTEDDGCFAVAST 169
 DB 117
 QY 165 GLITMTFSITELFTS 181
 DB 170 TELVLMFLTLAFCS 186

RESULT 6
 A42806
 IGE FC receptor beta chain - human
 C:Species: Homo sapiens (man)
 C>Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 08-Oct-1999
 C:Accession: A42806; S21154
 R:Kuester, H.; Zhang, L.; Brini, A.T.; MacGlashan, D.W.J.; Kinet, J.P.
 J. Biol. Chem. 267, 12782-12787, 1992
 A:Title: The gene and cDNA for the human high affinity immunoglobulin E receptor beta ch
 A:Reference number: A42806; MUID:92316966
 A:Accession: A42806
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-244 <RUE>
 A:Cross-references: GB:M89796; NID:g337417; PIDN:AAA60269.1; PID:g337418
 R:Maekawa, K.; Imagawa, N.; Tanaka, Y.; Harada, S.
 S Lett. 302, 161-165, 1992
 A:Title: Determination of the sequence coding for the beta subunit of the human high-aff
 A:Reference number: S21154; MUID:92339505
 C:Keywords: Immunoglobulin receptor; Transmembrane protein

Query Match 12.2%; Score 123.5; DB 2; Length 244;
 Best Local Similarity 25.6%; Pred. No. 0.00022;
 Matches 50; Conservative 33; Mismatches 93; Indels 19; Gaps 7;
 QY 9 PVLVPEPPETASEYESTELSATPTQSPLO-KLPARKMLIGIQLFGIMTFESFG 65
 DB 21 PABEVL--EISPEVSSGRLLKS--ASSPPLHWLVLYKKEQELVGTQTLIMICICFG 76
 QY 66 VIFLFTLLKPRPPPL-FLSGYFPGSVLFINGATLAVKRTTEFLILSRIMNL 123
 DB 77 TVVCASVLDSHIGDIFSSPKAGYFPGVGAIFFSISGMLSIISRRNATVIVRSLSGNTA 136

QY 124 SALRAINGIILTRGFILDONVICGYSHONS-----QCKAVTVLFLGILITLMTF-SILE 177
 DB 137 SSINGGTGTTL-----IINKSLAYIHHSCQKFEFFKCFMASFTEIVVMFLTLIG 192
 QY 178 LFLSPFSLGCHSE 192
 DB 193 LGSANVSLITGAGE 207

RESULT 7
 C86844
 hypothetical protein ysbC [imported] - Lactococcus lactis subsp. lactis (strain IL140
 C:Species: Lactococcus lactis subsp. lactis
 C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 23-Mar-2001
 C:Accession: C86844
 R:Bohloiti, A.; Winkler, P.; Mager, S.; Jallion, O.; Malarme, K.; Weissenbach, J.; Eh
 Genome Res. in press, 2001
 A:Title: The complete genome sequence of the lactic acid bacterium.
 A:Reference number: A86825
 A:Accession: C86844
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-307 <STO>
 A:Cross-references: GB:AE005176; NID:g12724776; PIDN:AAK05853.1; GSPDB:GN00146
 A:Experimental source: strain IL1403
 C:Genetics:
 A:Gene: ysbC

Query Match 10.1%; Score 103; DB 2; Length 307;
 Best Local Similarity 31.5%; Pred. No. 0.023;
 Matches 45; Conservative 22; Mismatches 50; Indels 26; Gaps 10;

QY 34 STQSPLOKLPARKKMLIGTQILFGIMTFESGVIFL--FTLLKPRPP-FFPLSGYFV 90
 DB 16 ANQPIN--ADLRKIVGSPFLASGINSF-VGSIFFGILITLVTQTLFSPQFVSHPAW 71
 QY 91 ---GSVFINSGAFLAVKRTTEFLILSRIMMLSLRAIAGIILTFEGFILDONVIC 147
 DB 72 IWIGGVV--GGIFL-----TSNVLLFPRLGAVQTVIPLIGQILM--GTLIDS--F 116
 QY 148 GYSHONSCKAVTVLFGILITL 170
 DB 117 GWFLH-WQLPMTLMRFLGIITL 138

RESULT 8
 JH0751
 IGE receptor beta chain - human (fragment)
 C:Species: Homo sapiens (man)
 C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
 C:Accession: JH0751
 R:Bieber, T.; de la Salle, H.; Wollenberg, A.; Hakim, J.; Chizzonite, R.; Ring, J.;
 J. Exp. Med. 175, 1285-1290, 1992
 A:Title: Human epidermal langerhans cells express the high affinity receptor for immu
 A:Reference number: JH0751; MUID:92235608
 A:Accession: JH0751
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-152 <BIE>
 A:Cross-references: GB:X66187; NID:g396463; PIDN:CAA6955.1; PID:g396464

Query Match 9.9%; Score 100; DB 2; Length 152;
 Best Local Similarity 27.6%; Pred. No. 0.023;
 Matches 32; Conservative 21; Mismatches 53; Indels 10; Gaps 3;
 QY 83 FLGYPFGSVLFINGATLAVKRTTEFLILSRIMMLSLRAIAGIILTFEGFILD 142
 DB 26 FKAGYFPGVGAIFFSISGMLSIISRRNATVIVRSLSGANTASSIAGGTGTTIL--IIN 81
 QY 143 QNVICGYSHONS-----QCKAVTVLFLGILITLMTF-SILELFLSPFSLGCHSE 192

Thu Oct 11 10:24:06 2001

us-09-735-712-2.rpr

Page 6

```

Db      344 TVSKAQTLIA-----LCGVPPWYETVFAFIDILGFEFTQMGFEFMTTEFMALAPYDHW 399
Oy      78  -----RFPPIFGSPWMSGVLFINSQAFILAVKRKRETELLIISRIINL-----ISALRAIA 130
Db      400 HRENKIGLIMVSLMFPAFNGPATITFVPAE-----IFPRALSTCHGISAASGRA 452
Oy      131 GILLTPEI-----LDQWYIGYSNONGQCAKAVVYLGLIGLITLM 171
Db      453 GAIVAGFEFLVRAOOSDSEKTDAGPPQIGYRNSLLMLACVNEIGVITILL 503

```

Search completed: October 7, 2001, 04:40:46
Job time: 1981 sec

QY 91 GSVLFNSGAFLLVKKRTETETIIISR-IMNLSALRAIAGIIL-----LTFGFIID 142
 DB 83 GGIMTIISSGLIAAAAKTKSRKSLVKKAVIMSSLSLFAAISGIIISLMDILNMLISFLK 142
 OY 143 QN-----YIC---GYSHONS---QCKAVYVLFGLITLTMTFSIIEFIS 181
 DB 143 MRRLIQTISKPYVDIYDCPSNSSEKNSPSTGYCNSIQSVFGLISLAMLISAFQRLVY 202

RESULT 2
 CD20_HUMAN STANDARD; PRT; 297 AA.
 ID CD20_HUMAN P11836; P08984; Q13963;
 AC P11836; P08984; Q13963;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE B-LYMPHOCYTE ANTIGEN CD20 (B-LYMPHOCYTE SURFACE ANTIGEN BI) (LEU-16) (BP35).
 MS4A2 OR CD20.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NC NCBL_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88258386; PubMed=3260267;
 RA Stamenkovic I., Seed B.;
 RT "Analysis of two cDNA clones encoding the B lymphocyte antigen CD20 (BI, Bp35), a type III integral membrane protein.";
 RT J. Exp. Med. 167:1973-1980(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88124792; PubMed=2448768;
 RA Tedder T.F., Streuli M., Schlossman S.F., Saito H.;
 RT "Isolation and structure of a cDNA encoding the BI (CD20) cell-surface antigen of human B lymphocytes.";
 RT Proc. Natl. Acad. Sci. U.S.A. 85:208-212(1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89176281; PubMed=2466899;
 RA Tedder T.F., Klejman G., Schlossman S.F., Saito H.;
 RT "Structure of the gene encoding the human B lymphocyte differentiation antigen CD20 (BI).";
 RT J. Immunol. 142:2560-2568(1989).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88283639; PubMed=2456210;
 RA Einfield D.A., Brown J.P., Valentine M.A., Clark E.A., Ledbetter J.A.;
 RT "Molecular cloning of the human B cell CD20 receptor predicts a hydrophobic protein with multiple transmembrane domains.";
 RT EMBO J. 7:711-717(1988).
 RL FUNCTION: THIS PROTEIN MAY BE INVOLVED IN THE REGULATION OF B-CELL ACTIVATION AND PROLIFERATION.
 CC SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- PTM: PHOSPHORYLATED. MIGHT BE FUNCTIONALLY REGULATED BY PROTEIN KINASE(S).
 CC -1- DATABASE: NAME=PROT; NOTE=CD guide CD20 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd20.htm".
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC EMBL: X12530; CAA31046.1; -
 CC EMBL: M27394; AAA35581.1; -
 CC EMBL: M27395; - NOT_ANNOTATED_CDS.
 CC EMBL: L23419; AAA88911.1; -
 CC EMBL: L23415; AAA88911.1; JOINED.

DR EMBL: L23416; AAA88911.1; JOINED.
 DR EMBL: L23417; AAA88911.1; JOINED.
 DR EMBL: X07203; CAA30178.1; -
 DR EMBL: X07204; CAA30180.1; -
 DR PIR: A27400; A27400.
 DR PIR: J10042; J10042.
 DR PIR: A30586; A30586.
 DR PIR: S00387; S00387.
 DR MIM: 112210; -
 KW B-cell; Transmembrane; Phosphorylation.
 FT DOMAIN 1 63 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 64 84 POTENTIAL.
 FT TRANSMEM 85 105 POTENTIAL.
 FT TRANSMEM 121 141 POTENTIAL.
 FT TRANSMEM 189 209 POTENTIAL.
 FT DOMAIN 210 297 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 81 167 PROBABLE.
 FT DISULFID 167 183 PROBABLE.
 FT DISULFID 111 220 PROBABLE.
 FT CONFLICT 13 13 P -> L (IN REF. 4).
 FT CONFLICT 71 71 M -> I (IN REF. 3).
 SQ SEQUENCE 297 AA; 33077 MW; AC5420F8B626BD1 CRC64;

Query Match 16.6%; Score 168; DB 1; Length 297;
 Best Local Similarity 30.3%; Pred. No. 9.6e-08;
 Matches 54; Conservative 27; Mismatches 63; Indels 34; Gaps 7;

QY 34 STQSPLOKLARKKILGTQILFGITFSFYIFLTLKPPRR-PIFFLSGYPPWGS 92
 DB 36 SLVGPQTSFFMRKSKTGLGAOVINGLPHIALGGI---IMIPGIVAPICVTWYPLMG 91
 QY 93 VLFNSGAFLLVKKRTETETIIISR-IMNLSALRAIAGIIL-----T 136
 DB 92 IMTIISSGLIAAAAKTKSRKSLVKKAVIMSSLSLFAAISGIIISLMDILNMLISFLK 151
 QY 137 FGFLLDN-YICG-----SHONS---QCKAVYVLFGLITLTMTFSII-ELFIS 181
 DB 152 LNFIRAHPTPIININCEPANPSEKNSPSTGYCNSIQSVFGLISLAMLISAFQRLV 209

RESULT 3
 FCER_RAT STANDARD; PRT; 243 AA.
 ID FCER_RAT P13386;
 AC P13386;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR BETA-SUBUNIT (FCERI) (IGE FC RECEPTOR, BETA-SUBUNIT).
 GN MS4A2 OR FCER1B OR FCER1B.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NC NCBL_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 11-29; 37-43 AND 217-243.
 RX MEDLINE=88320465; PubMed=2970642;
 RA Kinet J.-P., Blank U., Ra C., White K., Metzger H., Kochan J.;
 RT "Isolation and characterization of cDNAs coding for the beta subunit of the high-affinity receptor for immunoglobulin E.";
 RT Proc. Natl. Acad. Sci. U.S.A. 85:6483-6487(1988).
 RL FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULIN EPSILON. HIGH AFFINITY RECEPTOR. RESPONSIBLE FOR INITIATING THE ALLERGIC RESPONSE. BINDING OF ALLERGEN TO RECEPTOR-BOUND IGE LEADS TO CELL ACTIVATION AND THE RELEASE OF MEDIATORS (SUCH AS HISTAMINE) RESPONSIBLE FOR THE MANIFESTATIONS OF ALLERGY. THE SAME RECEPTOR ALSO INDUCES THE SECRETION OF IMPORTANT LYMPHOKINES.
 CC -1- SUBUNIT: TETRAMER OF AN ALPHA CHAIN, A BETA CHAIN, AND TWO DISULFIDE LINKED GAMMA CHAINS.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M22923; AAA41149.1; -
DR PIR: A31231; A31231.
KM IGE-binding protein; Receptor; Transmembrane.
FT DOMAIN 1 59 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 60 79 POTENTIAL.
FT DOMAIN 80 97 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 98 117 POTENTIAL.
FT DOMAIN 118 130 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 131 150 POTENTIAL.
FT DOMAIN 151 179 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 180 199 POTENTIAL.
FT DOMAIN 200 243 CYTOPLASMIC (POTENTIAL).
SO SEQUENCE 243 AA; 26730 MW; 471DFA59B688B5D CRC64;

Query Match 15.7%; Score 159; DB 1; Length 243;
Best Local Similarity 28.0%; Pred. No. 4.7e-07;
Matches 45; Conservative 29; Mismatches 71; Indels 16; Gaps 3;

QY 36 QSPLOKLFARKMKILGTLQIFGIMTFSGVIFLFLKRP--YPRPFIFLGGYPMGVS 93
DB 47 QQWQSLKLELEFLGTVGVGLICLFCSTVCSITLQTSDFDEVLILRAGIPFGAV 106
QY 94 LFNISGFLIAVRKKTETELLISRLINLLSALRAIAGIILTFGFIIDONYICGYSHQ 153
DB 107 LFLVSGFLISMERKNTLYVRSGLANIVSIAAGIAILILNLSNNAYM----- 159
QY 154 SOCKAVT-----VLFGLITLMTFSIIEFLISPLSL 187
DB 160 NYCDIEDDCGCVTSFTLVLMLFLFTILAFCSAVLLII 200

RESULT 4
FCER_MOUSE STANDARD; PRT; 235 AA.
AC P20490;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, last sequence update)
DT 01-OCT-2000 (Rel. 40, last annotation update)
DE HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR BETA-SUBUNIT (FCER1
DE (IGE FC RECEPTOR, BETA-SUBUNIT).
GN MS4A2 OR FCER1B OR FCER1B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN NCB1_TaxID=10090;
RP SEQUENCE FROM N.A.
RA MEDLINE=89359361; PubMed=2527850;
RA Ra C., Jouvin M.H.E., Kinet J.-P.;
RT "Complete structure of the mouse mast cell receptor for IGE (Fc
RT epsilon RI) and surface expression of chimeric receptors (rat-mouse-
RT human) on transfected cells";
RL J. Biol. Chem. 264:15323-15327(1989).
CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULIN EPSILON. HIGH
CC AFFINITY RECEPTOR. RESPONSIBLE FOR INITIATING THE ALLERGIC
CC RESPONSE. BINDING OF ALLERGEN TO RECEPTOR-BOUND IGE LEADS TO CELL
CC ACTIVATION AND THE RELEASE OF MEDIATORS (SUCH AS HISTAMINE)
CC RESPONSIBLE FOR THE MANIFESTATIONS OF ALLERGY. THE SAME RECEPTOR
CC ALSO INDUCES THE SECRETION OF IMPORTANT LYMPHOKINES.
CC -1- SUBUNIT: TETRAMER OF AN ALPHA CHAIN, A BETA CHAIN, AND TWO
CC DISULFIDE LINKED GAMMA CHAINS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: J05019; AAA37601.1; -
DR PIR: B34342; B34342.
DR MGD: MGI:95495; Fcer1b.
KM IGE-binding protein; Receptor; Transmembrane.
FT DOMAIN 1 51 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 52 71 POTENTIAL.
FT DOMAIN 72 89 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 90 109 POTENTIAL.
FT DOMAIN 110 122 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 123 142 POTENTIAL.
FT DOMAIN 143 171 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 172 191 POTENTIAL.
FT DOMAIN 192 235 CYTOPLASMIC (POTENTIAL).
SO SEQUENCE 235 AA; 25963 MW; 1C2D6BF195738168 CRC64;

Query Match 15.6%; Score 158.5; DB 1; Length 235;
Best Local Similarity 30.5%; Pred. No. 5.1e-07;
Matches 60; Conservative 29; Mismatches 81; Indels 27; Gaps 9;

QY 1 MDSSIAHSPFLVFP--PEIASESESTELSTFSTQSPQ--KLPAK-MKILGTLQIL 56
DB 1 MDREN-RSRADLALPNQESSAPDIELLSAPKAAPQYMTFLKLEFLGATQIL 59
QY 57 FGIMTFSGVIFLFLKRP-----PRPFIFLGGYPMGVSFLINSGFLIAVRKKTTE 111
DB 60 VGLICLFCSTVCSITLQTSDFDEVLILRAGIPFGAVLFLVSGFLISMERKNTL 116
QY 112 TLITLIRINMLLSALRAIAGIILTFGFIIDONYICGYSHQCKAVT-----VLF 164
DB 117 YLVRSGLANIVSIAAGIAILILNLSNNAYM-----NLTNNFAYMNN-CKNVTEDDCGCVASFT 169
QY 165 GILITLMTFSIIEFLIS 181
DB 170 TELVIMMLFLITLAFCS 186

RESULT 5
FCER_HUMAN STANDARD; PRT; 244 AA.
AC Q01362;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, last sequence update)
DT 01-OCT-2000 (Rel. 40, last annotation update)
DE HIGH AFFINITY IMMUNOGLOBULIN EPSILON RECEPTOR BETA-SUBUNIT (FCER1
DE (IGE FC RECEPTOR, BETA-SUBUNIT).
GN MS4A2 OR FCER1B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
RN NCB1_TaxID=9606;
RP SEQUENCE FROM N.A.
RA MEDLINE=92316966; PubMed=1535625;
RA Kuester H., Zhang L., Brini A.T., Macglaishan D.W., Kinet J.-P.;
RT "The gene and cDNA for the human high affinity immunoglobulin E
RT receptor beta chain and expression of the complete human receptor";
RL J. Biol. Chem. 267:12782-12787(1992).
CC [2]
CC SEQUENCE FROM N.A.
CC MEDLINE=92339505; PubMed=1386024;
CC Maekawa K., Imagawa N., Tanaka Y., Harada S.;
RT "Determination of the sequence coding for the beta subunit of the
RT human high-affinity IGE receptor";
RL FEBS Lett. 302:161-165(1992).
CC [3]
RN VARIANT G10-237.

RX MEDLINE-96414302; PubMed-8817330;
RA Hill M.R., Cookson W.O.;
RT "A new variant of the beta subunit of the high-affinity receptor for
immunoglobulin E (Fc epsilon RI-beta E237G): associations with
measures of atopy and bronchial hyper-responsiveness.";
RL Hum. Mol. Genet. 5:959-962(1996).
RN (4)
RP VARIANT GLU-237.
RX MEDLINE-96440420; PubMed-8842731;
RA Shirekawa T., Mao X.-Q., Sasaki S., Enomoto T., Kawai M., Morimoto K.,
RT Hopkin J.;
RT "Association between atopic asthma and a coding variant of Fc-epsilon-
RI-beta in a Japanese population.";
RL Hum. Mol. Genet. 5:1129-1130(1996).
RN (5)
RP ERRATUM.
RX MEDLINE-97123518; PubMed-8968765;
RA Shirekawa T., Mao X.-Q., Sasaki S., Enomoto T., Kawai M., Morimoto K.,
RT Hopkin J.;
RT Hum. Mol. Genet. 5:2068-2068(1996).
CC -1- FUNCTION: BINDS TO THE FC REGION OF IMMUNOGLOBULIN EPSILON. HIGH
AFFINITY RECEPTOR. RESPONSIBLE FOR INITIATING THE ALLERGIC
RESPONSE. BINDING OF ALLERGEN TO RECEPTOR-BOUND IGE LEADS TO CELL
ACTIVATION AND THE RELEASE OF MEDIATORS (SUCH AS HISTAMINE)
RESPONSIBLE FOR THE MANIFESTATIONS OF ALLERGY. THE SAME RECEPTOR
ALSO INDUCES THE SECRETION OF IMPORTANT LYMPHOKINES.
CC -1- SUBUNIT: TETRAMER OF AN ALPHA CHAIN, A BETA CHAIN, AND TWO
DISULFIDE LINKED GAMMA CHAINS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: FOUND ON THE SURFACE OF MAST CELLS AND
BASOPHILS.
CC -1- POLYMORPHISM: VARIANT GLU-237 HAS BEEN FOUND TO BE PRESENT IN ABOUT
5.3% OF A 1004 INDIVIDUALS POPULATION SAMPLE IN AUSTRALIA. IT
SEEMS TO BE A RISK FACTOR FOR ATOPIC DERMATITIS AND ASTHMA.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL: D10583; BAA01440.1; -
DR EMBL: M89796; AAA60269.1; -
DR PIR: S21154; S21154.
DR PIR: A42806; A42806.
DR MIM: 147138; -
CC IGE-binding protein; Receptor; Transmembrane; Polymorphism.
FT DOMAIN 1 59
FT TRANSMEM 60 79
FT DOMAIN 80 97
FT TRANSMEM 98 117
FT DOMAIN 118 130
FT TRANSMEM 131 150
FT DOMAIN 151 180
FT TRANSMEM 181 200
FT DOMAIN 201 244
FT TRANSMEM 245 263
FT VARIANT 264 283
FT SEQUENCE 284 283 /FTID=VAR_003965.
SQ SEQUENCE 244 AA; 26533 MW; CE52310ZD5F567AF CRC64;

Query Match 12.2%; Score 123.5; DB 1; Length 244;
Best Local Similarity 25.6%; Pred. No. 0.00056;
Matches 50; Conservative 33; Mismatches 93; Indels 19; Gaps 7;

OY 9 PVFLVPEETITASEYESTELATTFSTOSPLOQ--KLFARKMKILGTIOILFGIMTFSG 65
DB 21 PAREVL--EISPOVSSGRLLKS--ASSPLHMLVFLKKEQELFTYQILFLAMICCFG 76
OY 66 VIFLFLLLKRPFPFI--FLSGYFPGVSLFLNSGAFLLAVRKKTTEVITIIISRIIMNL 123

DB 77 TVVCSYVDISHIEDGIDJFSSEKAGYFPGWGAIFFSISGMLSIISERRNATYLVRSLSGANTA 136
OY 124 SALRAIGIILTFEGFLDNDNYICGVSHONS-----CKRAVYVFLGILTLMTF--SIIE 177
DB 137 SSIAGSGITIL-----IINKSLATYHHHSCKKFPETKCFMASFSFEIYVMFLFLITIG 192
OY 178 LFTSLPFSILGCHSE 192
DB 193 LGSASVSLTGAGEE 207

RESULT 6
TCR_STRAG STANDARD; PRT; 458 AA.
AC P13924;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE TETRACYCLINE RESISTANCE PROTEIN.
GN TEST.
OS Streptococcus agalactiae.
OC Plasmid pmv158.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1311;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE-90016790; PubMed-2677995;
RA van der Lelie D., Bron S., Venema G., Oskam L.;
RT "Similarity of minus origins of replication and flanking open reading
frames of plasmids pUB10, pTB913 and pmv158.";
RL Nucleic Acids Res. 17:7283-7294(1989).
CC -1- FUNCTION: RESISTANCE TO TETRACYCLINE BY AN ACTIVE TETRACYCLINE
EFFLUX. THIS IS AN ENERGY-DEPENDENT PROCESS THAT DECREASES THE
ACCUMULATION OF THE ANTIBIOTIC IN WHOLE CELLS. THIS PROTEIN
FUNCTIONS AS A METAL-TETRACYCLINE/H+ ANTI-PORTER.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
AS THE DRUG RESISTANCE TRANSLOCASE FAMILY).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL: X15669; CAA33712.1; -
DR PIR: C25599; YTSOG.
DR InterPro: IPR001411; -
DR PRINTS: PR01036; TCRPFETB.
KW Antibiotic resistance; Transmembrane; Transport; Symport; Plasmid.
FT TRANSMEM 12 33
FT TRANSMEM 81 100
FT TRANSMEM 111 129
FT TRANSMEM 140 162
FT TRANSMEM 165 185
FT TRANSMEM 201 221
FT TRANSMEM 223 240
FT TRANSMEM 256 276
FT TRANSMEM 297 317
FT TRANSMEM 324 344
FT TRANSMEM 346 365
FT TRANSMEM 432 451
FT SEQUENCE 458 AA; 50006 MW; AD2014E7CA199995 CRC64;

Query Match 9.1%; Score 92.5; DB 1; Length 458;
Best Local Similarity 25.3%; Pred. No. 0.49; Length 458;
Matches 44; Conservative 27; Mismatches 68; Indels 35; Gaps 7;

OY 32 TFSTOSPLOKLFARKMKILGTIOILFGIMTFSGVIFLFTLLKRPFPFIPLSGYFPG 91

```

DB 178 TITVPEMLMLKKEVRKIKHFDIK-GILMSGVIVF-----MLFTTSYI--- 223
QY 92 SVLFINSGAFLIAVK--RKTEFLILISRLMMLLSALRAIAGIILL--TFGFLIDQNYIC 147
DB 224 SFLIVSVLSLSEFLIEFKKIRKVTDFVDPGLCKNILEFMIGVLCGGIETFTVAGFSMPWYPM 283
QY 148 GYSHONSQCKAVNYFLP-----GILL-----TMTFSIIEFLISLP 184
DB 284 KDVAHOLSTAEIGSVILFPGTMSVILIEGYIGGILVDRGRGPLVILNTGVTLVSFS 337

RESULT 7
NUM_DROVA STANDARD: PRT; 324 AA.
AC P07710;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 1 (EC 1.6.5.3).
MT:ND1 OR ND1.
Drosophila yakuba (Fruit fly).
Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7245;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=2317.6 IVORY COAST;
RC MEDLINE=86089137; PubMed=3001325;
RA Clary D.O., Wolstenholme D.R.;
RT "The mitochondrial DNA molecular of Drosophila yakuba: nucleotide
RT sequence, gene organization and genetic code."
RT J. Mol. Evol. 22:252-271(1985).
RN [2];
RP SEQUENCE FROM N.A.
RC MEDLINE=84221393; PubMed=6328435;
RA Clary D.O., Wolstenholme D.R.;
RT "Sequence and arrangement of the genes for cytochrome b, URF1, URF4,
RT URF4, URF5, URF6 and five tRNAs in Drosophila mitochondrial DNA."
RT Nucleic Acids Res. 12:3747-3762(1984).
CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 1 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X03240; CA26997.1; -
CC PIR: D30020; D30020.
CC FLYBase; FBgn0013183; Dyak\mt:ND1.
CC InterPro: IPR001694; -
CC Pfam: PF00146; NADhdh: 1.
CC PROSITE: PS00667; COMPLEX1_ND1_1; 1.
CC PROSITE: PS00668; COMPLEX1_ND1_2; 1.
CC KMOxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.
CC SEQUENCE 324 AA; 37529 MW; D3F20A5968463196 CRC64;

Query Match 9.1%; Score 92; DB 1; Length 324;
Best local Similarity 26.3%; Pred. No. 0.39;
Matches 50; Conservative 23; Mismatches 63; Indels 54; Gaps 9;

QY 48 KILGTIOI-----LFGIMFSEFGVIFLTKPYRPP-----FIFISGY- 87
DB 30 KVLGYIQRKGPVKVGMGPFCALIKLFTKEQITPLISNLYSYISFISFLSLV 89
QY 88 ----PFWGSVLFIN-SGAFLIAVKRKTEFLILISRLM-----LSALRAIA----- 130

```

```

DB 90 WMCPFFVYKLSFNLGLFLCCSLGCVTYVMAGNSNSVYALGLBRVAQRTSYEVS 149
QY 131 -GIILLTFEFTL-DQNYICGYSHONSQCKAVNYFLICILITMTFSIIEFLISL----- 182
DB 150 LALIMLSFIFLIGSYNMVIFEYQ-----IYMFLLIFPMSIWMVILISIAETNRT 200
QY 183 PFSLIGCHSE 192
DB 201 PEDFAGESE 210

RESULT 8
NUM_CHOCR STANDARD: PRT; 666 AA.
AC P48920;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 5 (EC 1.6.5.3).
GN ND5 OR NAD5.
OS Chondrus crispus (Carraheen).
OG Mitochondrion.
OC Eukaryota; Rhodophyta; Florideophyceae; Gigartinales; Gigartinales;
OC Chondrus.
OX NCBI_TaxID=2769;
RN [1];
RP SEQUENCE FROM N.A.
RC TISSUE=Apices;
RC MEDLINE=95341681; PubMed=7616569;
RA Leblanc C., Boyen C., Richard O., Bonnard G., Grienenberger J.M.,
RA Kloareg B.;
RT "Complete sequence of the mitochondrial DNA of the rhodophyte
RT Chondrus crispus (Gigartinales). Gene content and genome
RT organization."
RT J. Mol. Biol. 250:484-495(1995).
RN [1];
RP -1- CATALYTIC ACTIVITY: NADH + ACCEPTOR = NAD(+) + REDUCED ACCEPTOR.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Z47547; CA87625.1; -
CC Mendel: 7732; CHOCR; nd5:1.
CC InterPro: IPR001516; -
CC InterPro: IPR001750; -
CC Pfam: PF00361; oxidored_q1; 1.
CC DR Pfam: PF00662; oxidored_q1_N; 1.
CC KMOxidoreductase; NAD; Ubiquinone; Mitochondrion.
CC SEQUENCE 666 AA; 75628 MW; 37E86F2C24B9D360 CRC64;

Query Match 9.0%; Score 91.5; DB 1; Length 666;
Best local Similarity 21.2%; Pred. No. 0.86;
Matches 56; Conservative 32; Mismatches 83; Indels 93; Gaps 9;

QY 16 PEITASEYESTELSAVTFSTOSPFLKLFARKMKILGTIOILFGIMFSEFGVIFLTKLP 75
DB 394 FLTGYSKDFLELFOISSYSLNLSQMSYISFACWLGTMGVFF-TSFYSFLIYTLFLNNT 452
QY 76 -----YRPEPFIFL-----SGYPPWGSVLFIN----- 98
DB 453 NLAQSSLNIVHSSSLMIFPLILISIGSIFAGYLIRDLFEVGSQSDPWGAIFILPKHSTP 512
QY 99 -----GAFLIAVKRKTEFLILISRLMMLLSALRAIAG----- 131
DB 513 IEAEELPIYWKMLPFLISLIGLFFASFPVQIFLKTFFKSNLQNLSEFFLILKKRWYMDV 572
QY 132 -----IILTFRGF-----IILDQNYI-----CGYSHONS-----QCKAVNYFLIG 165

```

Db 573 LYNNELVPLINFCYSSISFILDGFIELSGPYGFKVFSWQILIKITGOITHYLFF 632
 OY 166 ILITMTFSIIEI--FISLPEISIL 187
 Db 633 MIFFCFSIILVSYINILFNIL 656

RESULT 9

NUIM_DROSU STANDARD; PRT; 312 AA.
 ID NUIM_DROSU
 AC P51937;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 1 (EC 1.6.5.3).
 GN MT:ND1 OR ND1.
 OS Drosophila subobscura (Fruit fly).
 Mitochondrion.
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7241;
 RN [1]
 RP SEQUENCE OF 1-152 FROM N.A.
 RX MEDLINE=95106285; PubMed=7528808;
 RA Barrio E., Latorre A., Moya A.;
 RT "Phylogeny of the Drosophila obscura species group deduced from
 mitochondrial DNA sequences."
 RL J. Mol. Evol. 39:478-486(1994).
 RN [2]
 RP SEQUENCE OF 50-312 FROM N.A.
 RA STRAIN=SSP. TUE 3;
 RX Volz-Lingenhl A.;
 RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
 CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 1 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U07288; AAA6626.1; -
 DR EMBL: X65130; CAA46260.1; -
 DR FlyBase: FBgn0012956; Dsubmnt:ND1.
 DR InterPro: IPR001694; -
 DR Pfam: PF00146; NADHdh; 1.
 DR PROSITE: PS00667; COMPLEX1_ND1_1; 1.
 DR PROSITE: PS00668; COMPLEX1_ND1_2; 1.
 KW Mitochondrion; Oxidoreductase; NAD; Transmembrane.
 SO SEQUENCE 312 AA; 35877 MW; AFD1B33BC2BCDE68 CRC64;

Query Match 8.9%; Score 90; DB 1; Length 312;
 Best Local Similarity 25.8%; Pred. No. 0.56;
 Matches 49; Conservative 24; Mismatches 63; Indels 54; Gaps 9;

OY 48 KILGTIOI-----LFGIMTFSFGVIFLFTLKPPRP-----FIFLSGY- 87
 Db 30 KVLGYQIRKPNKVGIMGPQPCDAIKLFTKEQYPLISNLYSISPIFSLFLV 89
 OY 88 -----PFGMSVLFIN--SGAFLIAVKRKTTETLLISRLM-----LSALRAIA----- 130
 Db 90 WMCMPFVKLIYSNGLGFLFCLCTSLGYTVVAVAGSSSNVALLGGLRAVAQTISYEVS 149
 OY 131 -GILLTFEFLI--DQNYICGYSHONSQCAVAVLFGILITMTFSIIEFISL----- 182
 Db 150 LAILLSFVLISGYNNVVFYQ-----IYWFLLILFPALVWLTISLAETNRT 200

OY 183 PFSILGCHSE 192
 Db 201 PFDFAEGESE 210

RESULT 10

NUIM_DROME STANDARD; PRT; 312 AA.
 ID NUIM_DROME
 AC P18929;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 1 (EC 1.6.5.3).
 GN MT:ND1 OR ND1.
 OS Drosophila melanogaster (Fruit fly).
 Mitochondrion.
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRETAGNE;
 RX MEDLINE=88212147; PubMed=3130291;
 RA Garesse R.;
 RT "Drosophila melanogaster mitochondrial DNA: gene organization and
 RT evolutionary considerations."
 RL Genetics 118:649-663(1988).
 CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
 CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 1 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M37275; AAA69715.1; -
 DR EMBL: U37541; AAC47823.1; -
 DR PIR: S01191; S01191.
 DR FlyBase: FBgn0013679; mt:ND1.
 DR InterPro: IPR001694; -
 DR Pfam: PF00146; NADHdh; 1.
 DR PROSITE: PS00667; COMPLEX1_ND1_1; 1.
 DR PROSITE: PS00668; COMPLEX1_ND1_2; 1.
 KW Oxidoreductase; NAD; ubiquinone; Mitochondrion; Transmembrane.
 SO SEQUENCE 312 AA; 35910 MW; B2B01BE6213CE722 CRC64;

Query Match 8.6%; Score 87; DB 1; Length 312;
 Best Local Similarity 25.8%; Pred. No. 1;
 Matches 49; Conservative 24; Mismatches 63; Indels 54; Gaps 9;

OY 48 KILGTIOI-----LFGIMTFSFGVIFLFTLKPPRP-----FIFLSGY- 87
 Db 30 KVLGYQIRKPNKVGIMGPQPCDAIKLFTKEQYPLISNLYSISPIFSLFLV 89
 OY 88 -----PFGMSVLFIN--SGAFLIAVKRKTTETLLISRLM-----LSALRAIA----- 130
 Db 90 WMCMPFVKLIYSNGLGFLFCLCTSLGYTVVAVAGSSSNVALLGGLRAVAQTISYEVS 149
 OY 131 -GILLTFEFLI--DQNYICGYSHONSQCAVAVLFGILITMTFSIIEFISL----- 182
 Db 150 LAILLSFVLISGYNNVVFYQ-----IYWFLLILFPALVWLTISLAETNRT 200
 OY 183 PFSILGCHSE 192
 Db 201 PFDFAEGESE 210

RESULT 11

RFC_SHIFL
ID RFC_SHIFL STANDARD: PRT: 382 AA.
AC P37784;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE O-ANTIGEN POLYMERASE.
GN RFC.
OS Shigella flexneri.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SEROTYPE 2A;
RX MEDLINE=94131953; PubMed=7507920;
RA Morona R., Mavris M., Fallarino A., Manning P.A.;
RT "Characterization of the rfc region of Shigella flexneri";
RT J. Bacteriol. 176:733-747(1994).
CC -1- FUNCTION: MAY LINK THE O-ANTIGEN TETRASACCHARIDE UNITS INTO LONG
CHAINS, GIVING RISE TO TYPICAL SMOOTH LPS.
CC -1- PATHWAY: LIPOLYSACCHARIDE BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
(POTENTIAL).
CC -1- SIMILARITY: TO OTHER O-ANTIGEN POLYMERASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X11970; CAA50774.1;
KW Lipopolysaccharide biosynthesis; Transmembrane; Inner membrane.
FT TRANSMEM 11 31 POTENTIAL.
FT TRANSMEM 36 56 POTENTIAL.
FT TRANSMEM 61 81 POTENTIAL.
FT TRANSMEM 87 107 POTENTIAL.
FT TRANSMEM 126 146 POTENTIAL.
FT TRANSMEM 166 186 POTENTIAL.
FT TRANSMEM 194 214 POTENTIAL.
FT TRANSMEM 228 248 POTENTIAL.
FT TRANSMEM 303 323 POTENTIAL.
FT TRANSMEM 325 345 POTENTIAL.
FT TRANSMEM 348 368 POTENTIAL.
SQ SEQUENCE 382 AA; 43723 MW; DCE02BB41619156A CRC64;

Query Match 8.6%; Score 87; DB 1; Length 382;
Best Local Similarity 19.7%; Pred. No. 1.2;
Matches 46; Conservative 39; Mismatches 64; Indels 84; Gaps 8;

ID TCR_BACST STANDARD: PRT: 458 AA.
AC P07561;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE TETRACYCLINE RESISTANCE PROTEIN.
GN TET.
OS Bacillus stearohermophilus.
OC plasmid pTHR15.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86031344; PubMed=2996983;
RA Hoshino T., Ikeda T., Tomizuka N., Furukawa K.;
RT "Nucleotide sequence of the tetracycline resistance gene of pTHR15, a
RT thermophilic Bacillus plasmid: comparison with staphylococcal tcr
RT controls.";
RT Gene 37:131-138(1985).
CC -1- FUNCTION: RESISTANCE TO TETRACYCLINE BY AN ACTIVE TETRACYCLINE
EFFLUX. THIS IS AN ENERGY-DEPENDENT PROCESS THAT DECREASES THE
CC ACCUMULATION OF THE ANTIBIOTIC IN WHOLE CELLS. THIS PROTEIN
CC FUNCTIONS AS A METAL-TETRACYCLINE/H+ ANTI-PORTER.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
CC AS THE DRUG RESISTANCE TRANSLOCASE FAMILY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M1036; AAA22851.1;
DR PIR: A23973; YTBSTR.
DR InterPro: IPR001411;
DR PRINTS: PRO1036; TCRTEB.
KW Antibiotic resistance; Transmembrane; Transport; Symport; Plasmid.
FT TRANSMEM 12 33 POTENTIAL.
FT TRANSMEM 81 100 POTENTIAL.
FT TRANSMEM 111 129 POTENTIAL.
FT TRANSMEM 140 162 POTENTIAL.
FT TRANSMEM 165 185 POTENTIAL.
FT TRANSMEM 201 221 POTENTIAL.
FT TRANSMEM 223 240 POTENTIAL.
FT TRANSMEM 256 276 POTENTIAL.
FT TRANSMEM 297 317 POTENTIAL.
FT TRANSMEM 324 344 POTENTIAL.
FT TRANSMEM 346 365 POTENTIAL.
FT TRANSMEM 432 451 POTENTIAL.
SQ SEQUENCE 458 AA; 50119 MW; E1F3F4C57E7B06A CRC64;

Query Match 8.4%; Score 85.5; DB 1; Length 458;
Best Local Similarity 24.7%; Pred. No. 2;
Matches 43; Conservative 27; Mismatches 69; Indels 35; Gaps 7;

RESULT 12
TCR_BACST

```

RESULT 13
TCR_STRPN STANDARD: PRT: 458 AA.
ID TCR_STRPN STANDARD: PRT: 458 AA.
AC P11063; P72219;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE TETRACYCLINE RESISTANCE PROTEIN.
GN TET.
OS Streptococcus pneumoniae, Bacillus cereus, and Bacillus subtilis.
OG Plasmid pLS1, Plasmid pBC16, plasmid pHY163PLK, Plasmid pTB19, and
OG Plasmid pMS1981.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus
OX NCBI_TaxID=1313, 1396, 1423;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.pneumoniae; PLASMID=PLS1;
RC MEDLINE=87226167; PubMed=2438417;
RA Lacks S.A., Lopez P., Greenberg B., Espinosa M.;
RA "Identification and analysis of genes for tetracycline resistance and
RA replication functions in the broad-host-range plasmid pLS1.";
RA J. Mol. Biol. 192:753-765(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=B.cereus; PLASMID=PBC16;
RC MEDLINE=90221899; PubMed=2109312;
RA Palva A., Vlodavet G., Simonen M., Rintala H., Laamanen P.;
RA "Nucleotide sequence of the tetracycline resistance gene of pBC16
RA from Bacillus cereus.";
RA Nucleic Acids Res. 18:1635-1635(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC PLASMID=pHY163PLK;
RA Ishiwa H., Shibaara H.;
RA "New shuttle vectors for Escherichia coli and Bacillus subtilis. III.
RA Nucleotide sequence analysis of tetracycline resistance gene of
RA pAMAlphal and ori-177.";
RA Jpn. J. Genet. 60:485-498(1985).
RN [4]
RP SEQUENCE FROM N.A.
RC PLASMID=pHY163PLK;
RC MEDLINE=83129391; PubMed=6186390;
RA Seizer G., Som T., Itoh T., Tomizawa J.;
RA "The origin of replication of plasmid p15A and comparative studies on
RA the nucleotide sequences around the origin of related plasmids.";
RA Cell 32:119-129(1983).
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES=B.subtilis; PLASMID=pTB19;
RC MEDLINE=92052681; PubMed=1946749;
RA Oskam L., Hillenga D.J., Venema G., Bron S.;
RA "The large Bacillus plasmid pTB19 contains two integrated
RA rolling-circle plasmids carrying mobilization functions.";
RA plasmid 26:30-39(1991).
RN [6]
RP SEQUENCE FROM N.A.
RC SPECIES=B.subtilis; STRAIN=168 / MARBURG; PLASMID=pMS1981;
RC MEDLINE=86287699; PubMed=3090576;
RA Sakaguchi R., Shishido K., Hoshino T., Furukawa K.;
RA "The nucleotide sequence of the tetracycline resistance gene of
RA plasmid pMS1981 from Bacillus subtilis differs from pTH15 from a
RA Thermophilic Bacillus by two base pairs.";
RA plasmid 16:72-73(1986).
RN [7]
RP FUNCTION: RESISTANCE TO TETRACYCLINE BY AN ACTIVE TETRACYCLINE
RN EFFLUX. THIS IS AN ENERGY-DEPENDENT PROCESS THAT DECREASES THE
RN ACCUMULATION OF THE ANTIBIOTIC IN WHOLE CELLS. THIS PROTEIN
RN FUNCTIONS AS A METAL-TETRACYCLINE/H+ ANTI-PORTER.
RN -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
RN -1- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
RN AS THE DRUG RESISTANCE TRANSPOCASE FAMILY).
RN -----
RN THIS SWISS-PROT entry is copyright It is produced through a collaboration

```

```

CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X51366; CAA35751.1; -
DR      EMBL; M29725; AAA98167.1; -
DR      EMBL; D13792; -; NOT_ANNOTATED_CDS.
DR      EMBL; M63891; AAA98304.1; -
DR      EMBL; D000006; BAA00005.1; -
DR      PIR; S09234; YTB5U6.
DR      PIR; G25599; YTS0G.
DR      PIR; J01211; J01211.
DR      InterPro; IPR001411. -
DR      PRINTS; PR01036; TCTETB.
DR      Antibiocic resistance; Transmembrane; Transport; Symport; Plasmid.
FT      TRANSMM      12      33      POTENTIAL.
FT      TRANSMM      81      100      POTENTIAL.
FT      TRANSMM      111      129      POTENTIAL.
FT      TRANSMM      140      162      POTENTIAL.
FT      TRANSMM      165      185      POTENTIAL.
FT      TRANSMM      201      221      POTENTIAL.
FT      TRANSMM      223      240      POTENTIAL.
FT      TRANSMM      256      276      POTENTIAL.
FT      TRANSMM      297      317      POTENTIAL.
FT      TRANSMM      324      344      POTENTIAL.
FT      TRANSMM      346      365      POTENTIAL.
FT      TRANSMM      432      451      POTENTIAL.
SO      SEQUENCE      458 AA; 50092 MW; 6A708777D44C2074 CRC64;

```

```

Query Match Similarity      8.4%; Score 85.5; DB 1; Length 458;
Best Local Similarity      24.7%; Pred. No. 2;
Matches 43; Conservative 27; Mismatches 69; Indels 35; Gaps
7;

OY 32 TESTOSPOKLFARMKILGTIQLIFGIMTFSGVIFLFTLLKPYRPPRFILSGYPMWG 91
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 178 TIITYPFLMKLLKREKVRILKGHFDIK-GIILMSVGIYF-----MFTTSYST-- 223
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY 92 SVLFINSGAFLLAVK-RKTTETLLISRLMNLSSALRAIGIIL--TFGFIIDONYIC 147
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 224 SFLIYSVSLFLEFVHKIKRVTDPFVDPDLGNRIPNIMICVLGGIIFGVAAFSVNPYMM 283
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

OY 148 GYSHNSOCCAKVTYFL-----GILI---TMTFSITELFISLPF 184
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 284 KDVHQLSTAEIGSVIIFPGTMSVLIIFGYGILVDRGRGLYVLNIGVFLSVSF 337
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 14
NU4M_DROME STANDARD; PRT: 446 AA.
AC PI8931;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE NADH-OBIDQUINONE OXIDOREDUCTASE CHAIN 4 (EC 1.6.5.3).
MT:ND4 OR ND4.
OS Drosophila melanogaster (Fruit fly).
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRETAGNE;
RX MEDLINE=88212147; PubMed=3130291;
RA Garesse R.;
RT "Drosophila melanogaster mitochondrial DNA: gene organization and
  evolutionary considerations.";
  Genetics 118:649-663(1988).
  TL

```


:Accession: A30558

QY 34 STQSPLOKLFARKMKILGTLQILFGIMTFSPGVFLFTLLKPYDRF-PFIFLSGPFMG 92

DB 36 SLVGPITQSFMRKESKTLGAVOIMNGLFHIALGL-----LMPAIVAPICVTWYPIIMG 91

QY 93 VLFINSQAFILVARKKTEETLGLITITMFSIIEFLISLPFSL 136

DB 92 IMYIISGSLAATEKNSRKL--VKGMIMNSISLFAISGMIL 133

RESULT 6
T37205
hypothetical protein Y57G7A.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T37205

R:Ozersky, P.
Submitted to the EMBL Data Library, March 1999

A:Description: The sequence of C. elegans cosmid Y57G7A.

A:Reference number: Z21634

A:Accession: T37205

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-457 <OE>

A:Cross-references: EMBL:AF07542; PIDN:AAC26295.1

C:Genetics:

A:Map position: II

A:Introns: 100/2; 193/2; 234/3; 272/2; 313/2; 399/1

A:Note: Y57G7A.4

Query Match 13.0%; Score 100; DB 2; Length 457;
Best Local Similarity 23.0%; Pred. No. 0.017;
Matches 41; Conservative 21; Mismatches 36; Indels 80; Gaps 6;

QY 14 FPPETITASESTELSATFTSTQSPLOKLFARKMKILGTLQILFGIMTFSPGVFLFTLL 73

DB 197 FPPVTVSEYMTWAKAYSKADQPLTLVFT-----ILDGLSKIIIPAVI----- 240

QY 74 KPYRPFPIFLSGVFPWGSVLFINSQAFILVARKKTEETLG-----LITL 119

DB 241 --YPIFTF-----LLIQQLRAAIALRKRTSTSGSRLESTKSDQTKMVTLM 286

QY 120 MTFESI-----LPISTPFI-----LQCHSE 141

DB 287 VTFISGPIGICYILEGTLPKRSVFEDINVDLMDAFTIVAINASVHFLICGVHSQ 344

RESULT 7
A42806
Fc receptor beta chain - human
C:Species: Homo sapiens (man)
C:Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 08-Oct-1999

C:Accession: A42806; S21154

R:Kuester, H.; Zhang, L.; Brini, A.T.; Macgashan, D.W.J.; Kinet, J.P.

J. Biol. Chem. 267, 12782-12787, 1992

A:Title: The gene and cDNA for the human high affinity immunoglobulin E receptor beta chain

A:Reference number: A42806; MIMID:92339506

A:Accession: A42806

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-244 <OE>

A:Cross-references: GB:M89796; NID:9337417; PIDN:AAA60269.1; PID:9337418

R:Maekawa, K.; Imagawa, N.; Tanaka, Y.; Harada, S.

FEBS Lett. 302, 161-165, 1992

A:Title: Determination of the sequence coding for the beta subunit of the human high-affinity

A:Reference number: S21154; MIMID:92339505

A:Accession: S21154

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-244 <MAE>

A:Cross-references: GB:D10583; NID:9219881; PIDN:BA01440.1; PID:d1001914; PID:9219882

C:Genetics:

A:Introns: 19/2; 62/3; 107/3; 126/3; 179/3; 212/3

C:Keywords: immunoglobulin receptor; transmembrane protein

Query Match 12.9%; Score 99.5; DB 2; Length 244;

Best Local Similarity 25.8%; Pred. No. 0.01;

Matches 40; Conservative 20; Mismatches 64; Indels 31; Gaps 6;

QY 9 PVFIVFPETITASESTELSATFTSTQSPLOQ---KLFARKMKILGTLQILFGIMTFSPG 65

DB 21 PAFEVLL--EISPOEVSSGRLLKS--ASSPLHTWLTVLKKEQFELGVQIITAMICLCFG 76

QY 66 VIFELTLLKPYRPFET--FLSGVFPWGSVLFINSQAFILVARKKTEETL----- 113

DB 77 TVQSVADISHIEDIDISSRKAGYFPFGAIFFSISGMLSISERNRNTYLVRSGLGNTA 136

QY 114 -----GLITLMTFSIIEFLISLPFSLGCHSEDC 143

DB 137 SSLAGGTGTYT-----LTIINLKSLAYIHHS 164

RESULT 8

T16070
hypothetical protein F14B8.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C:Accession: T16070

R:Geisel, C.
Submitted to the EMBL Data Library, June 1995

A:Description: The sequence of C. elegans cosmid F14B8.

A:Reference number: Z18456

A:Accession: T16070

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-455 <GEI>

A:Cross-references: EMBL:U028737; NID:9860717; PID:9860719; PIDN:AAA68272.1; CESP:F14B

A:Experimental source: strain Bristol N2

C:Genetics:

A:Gene: CESP:F14B8.3

A:Introns: 56/3; 193/2; 224/3; 269/2; 302/3; 370/2; 402/3; 442/3

Query Match 11.9%; Score 92; DB 2; Length 455;
Best Local Similarity 29.7%; Pred. No. 0.11;
Matches 27; Conservative 15; Mismatches 29; Indels 20; Gaps 3;

QY 49 ILGTQILFGIMT-----FSEGYIFLFTLLKPYRFPFIFLSGYFPWG 91

DB 303 LIGLVEIYFGLISRRMLHLISLCINGISMCASGVLLIFGLQKNPRTYLVFLIGY---A 359

QY 92 SVLFINSQAFILVARKKTEETLGLITLMTF 122

DB 360 STEFINDACTYLCVEQVPTVRGTISGACSF 390

RESULT 9
T02553
cellulose synthase homolog T26B15.10 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001

C:Accession: T02553; D84734

R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Syke

submitted to the EMBL Data Library, July 1998

A:Description: Arabidopsis thaliana chromosome II BAC T26B15 genomic sequence.

A:Reference number: Z14678

A:Accession: T02553

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-755 <ROU>

A:Cross-references: EMBL:AC004681; NID:93298532; PID:93298542

A:Experimental source: cultivar Columbia

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.

M.; Koo, H.; Moffatt, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,

ews, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2001, 04:46:09 ; Search time 24.63 Seconds

(Without alignments)
207.230 Million cell updates/sec

Title: US-09-735-712-8

Perfect score: 771
Sequence: 1 MDSSTAHSPVLPVFPPEITA.....SLPESILCHSEDCDCQCC 149

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	124	16.1	291	1	CD20_MOUSE
2	123.5	16.0	243	1	FCBE_RAT
3	117.5	15.2	235	1	FCBE_MOUSE
4	110.5	14.3	297	1	CD20_HUMAN
5	99.5	12.9	244	1	FCBE_HUMAN
6	82	10.6	657	1	NU5M_EMENT
7	81.5	10.6	590	1	CAN1_YEAST
8	80	10.4	458	1	TCR_BACST
9	80	10.4	458	1	TCR_STRAG
10	80	10.4	458	1	TCR_STRPN
11	78.5	10.2	573	1	ALP1_YEAST
12	77.5	10.1	497	1	ANSP_YEAST
13	77.5	10.0	669	1	COTE_HUMAN
14	75.5	9.8	669	1	NU5M_ARATH
15	75.5	9.8	803	1	ATCU_BACSU
16	74.5	9.7	641	1	NU5M_ALMA
17	74.5	9.7	652	1	NU5M_PODAN
18	74	9.6	402	1	YCH_BACSU
19	74	9.6	605	1	NU5M_CHICK
20	73.5	9.5	522	1	YOM3_CAEEL
21	73	9.5	439	1	LNT_AQUAE
22	72.5	9.4	438	1	CKR4_HUMAN
23	72.5	9.4	438	1	MAE1_SCHPO
24	72	9.3	458	1	TCR_STRAY
25	72	9.3	499	1	ANSP_ECOLI
26	72	9.3	664	1	NU5M_PHYIN
27	71.5	9.3	356	1	EMRD_ECOLI
28	71.5	9.3	459	1	NU4M_MOUSE
29	71	9.2	298	1	NU1M_ARTSF
30	71	9.2	315	1	PSTC_HAEIN
31	71	9.2	373	1	CKR2_RAT
32	71	9.2	562	1	FLA1_METUA
33	71	9.2	606	1	NU5M_EOUAS

34	71	9.2	854	1	VP2_BOVIN	097681 bos taurus
35	70.5	9.1	166	1	YC4B_METUA	P81230 methanococ
36	70.5	9.1	349	1	FM2_MACMU	P79191 macaca mula
37	70.5	9.1	460	1	NU4M_PIG	079881 sus scrofa
38	70.5	9.1	527	1	NU2M_ACACA	037376 acanthamoeb
39	70	9.1	436	1	SECY_METUA	060175 methanococ
40	70	9.1	604	1	NU5M_HORSE	P48656 equus cabal
41	70	9.1	606	1	NU5M_SHEEP	078756 ovis aries
42	69.5	9.0	121	1	YG18_YEAST	P53151 saccharomyc
43	69	8.9	175	1	Y433_METUA	057875 methanococ
44	69	8.9	458	1	TCR_BACSU	P23054 bacillus su
45	68.5	8.9	334	1	SRB7_CAEEL	P54142 caenorhabdi

ALIGNMENTS

RESULT	ID	CD20_MOUSE	STANDARD	PRT	291 AA
AC	1	CD20_MOUSE			
AC	1	CD20_MOUSE			
DT	01-FEB-1991 (Rel. 17, Created)				
DT	01-FEB-1991 (Rel. 17, Last sequence update)				
DT	01-OCT-2000 (Rel. 40, Last annotation update)				
DE	B-CELL SURFACE PROTEIN CD20 HOMOLOG (B-CELL DIFFERENTIATION ANTIGEN LY-44).				
GN	MSA42 OR CD20 OR LY-44.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
OX	NCBI_TaxID=10090;				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=89067519; PubMed=2461992;				
RA	Tedder T.F., Kiehlman G., Distche C.M., Adler D.A., Schlossman S.F.,				
RA	Saito H.;				
RT	"Cloning of a complementary DNA encoding a new mouse B lymphocyte				
RT	differentiation antigen, homologous to the human B1 (CD20) antigen,				
RT	and localization of the gene to chromosome 19."				
RL	J. Immunol. 141:4388-4394(1988).				
CC	-1- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN THE REGULATION OF				
CC	B-CELL ACTIVATION AND PROLIFERATION.				
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.				
CC	-1- PTM: PHOSPHORYLATED (BY SIMILARITY).				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement. (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	EMBL: M62541; AAA37394.1; -				
DR	PIR: A30558; A30558.				
DR	MGI: 88321; G320.				
KM	B-cell; Transmembrane; Phosphorylation.				
FT	DOMAIN 1 44				
FT	TRANSMEM 45 65				
FT	TRANSMEM 69 89				
FT	TRANSMEM 112 132				
FT	TRANSMEM 183 203				
FT	DOMAIN 204 291				
SQ	SEQUENCE 291 AA; 31958 MW; DF478ECD2C5C16FC CRC64;				

Query Match 16.1%; Score 124; DB 1; Length 291;

Best local Similarity 33.0%; Pred. No. 0.00017;

Matches 35; Conservative 18; Mismatches 47; Indels 6; Gaps 3;

DB 32 TSTSTSPLOKLFARKKMLGTOILFGIMTFSGVIFLTLKPYRFPFIFLS-GYFPW 90
27 TSTLVPTGPTGFFMRKSKALGAVOIMNGLPHITLGL- - - - -LMIPGVFAFICLSVYPLW 82


```

DE 01-NOV-1988 (Rel. 09, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE B-LYMPHOCYTE ANTIGEN CD20 (B-LYMPHOCYTE SURFACE ANTIGEN B1) (LEU-16)
DE (Bp35).
GN MS4A2 OR CD20.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8828386; PubMed=3260267;
RA Stamenkovic I., Seed B.;
RT "Analysis of two cDNA clones encoding the B lymphocyte antigen CD20
RT (B1, Bp35), a type III integral membrane protein.";
RL J. Exp. Med. 167:1975-1980(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88124792; PubMed=2448768;
RA Tedder T.F., Streuli M., Schlossman S.F., Saito H.;
RT "Isolation and structure of a cDNA encoding the B1 (CD20) cell-surface
RT antigen of human B lymphocytes.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:208-212(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=89176281; PubMed=2466899;
RA Tedder T.F., Klejman G., Schlossman S.F., Saito H.;
RT "Structure of the gene encoding the human B lymphocyte
RT differentiation antigen CD20 (B1).";
RL J. Immunol. 142:2560-2568(1989).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=88283639; PubMed=2456210;
RA Elfried D.A., Brown J.P., Valentine M.A., Clark E.A., Iedebetter J.A.;
RT "Molecular cloning of the human B cell CD20 receptor predicts a
RT hydrophobic protein with multiple transmembrane domains.";
RL EMBO J. 7:711-717(1988).
CC -I- FUNCTION: THIS PROTEIN MAY BE INVOLVED IN THE REGULATION OF
CC B-CELL ACTIVATION AND PROLIFERATION.
CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -I- PTM: PHOSPHORYLATED. MIGHT BE FUNCTIONALLY REGULATED BY PROTEIN
CC KINASE(S).
CC -I- DATABASE: NAME=PROV; NOTE=CD guide CD20 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd20.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC DR EMBL, X12530; CAA31046.1; -.
CC DR EMBL, M27394; AAA35581.1; -.
CC DR EMBL, M27395; -. NOT_ANNOTATED_CDS.
CC DR EMBL, L23419; AAA88911.1; -.
CC DR EMBL, L23415; AAA88911.1; JOINED.
CC DR EMBL, L23416; AAA88911.1; JOINED.
CC DR EMBL, L23417; AAA88911.1; JOINED.
CC DR EMBL, X07203; CAA30179.1; -.
CC DR EMBL, X07204; CAA30180.1; -.
CC DR PIR, A27400; A27400.
CC DR PIR, J10042; J10042.
CC DR PIR, A30586; A30586.
CC DR PIR, S00387; S00387.
CC DR MIM, 112210; -.
CC
CC FM B-cell; Transmembrane; Phosphorylation.
CC FT DOMAIN 1 63 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 64 84 POTENTIAL.
CC FT TRANSMEM 85 105 POTENTIAL.
CC FT TRANSMEM 121 141 POTENTIAL.

```

FT	TRANSMEM	189	209	POTENTIAL.
FT	DOMAIN	210	297	CYTOPLASMIC (POTENTIAL).
FT	DISULFID	81	167	PROBABLE.
FT	DISULFID	167	183	PROBABLE.
FT	DISULFID	111	220	PROBABLE.
FT	CONFLICT	13	13	P -> I (IN REF. 4).
FT	CONFLICT	71	71	M -> I (IN REF. 3).
SO	SEQUENCE	297 AA;	33077 MW;	AC5420F8B626BDD1 CRC64;

Query Match	14.3%;	Score 110.5;	DB 1;	Length 297;
Best Local Similarity	28.8%;	Pred. No. 0.0029;		
Matches 30;	Conservative 21;	Mismatches 46;	Indels 7;	Gaps

Qy	34	STOSPLQKFAKKKIKIGTQILFGLMTSEFGYIFLTLLKPYRP-PRIFLSGYPFMG	92
Db	36	SLVGPTQSEFNRBSKTLGAVQIMNGLEHIALGL-----LMPIGIAPICVWVWPLMG	91
Qy	93	VLFINSGAFLIAVKRKTTEFTLITLMPFSITELFISLPSFL	136
Db	92	IMVYISGLLAATEKNSRKL-VKGMWNSLSPAISGMIL	133

RESULT 5	FCERB_HUMAN	STANDARD;	PRT;	244 AA.
AC	Q01362;			
DT	01-APR-1993 (Rel. 25, Created)			
DT	01-APR-1993 (Rel. 25, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	HIGH AFFINITY IMMUGLOBULIN EPSILON RECEPTOR BETA-SUBUNIT (FCERI)			
DE	(10E FC RECEPTOR, BETA-SUBUNIT).			
GN	MS4A2 OR FCER1B.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92316966; PubMed=1535625;			
RA	Kuester H., Zhang L., Brini A.T., Macglaashan D.W., Kinet J.-P.;			
RT	"The gene and cDNA for the human high affinity immunoglobulin E			
RL	J. Biol. Chem. 267:12782-12787(1992).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92339505; PubMed=1366024;			
RA	Meekawa K., Imagawa N., Tanaka Y., Harada S.;			
RT	"Determination of the sequence coding for the beta subunit of the			
RL	human high-affinity IGE receptor.";			
RP	PEBS Lett. 302:161-165(1992).			
RN	[3]			
RP	VARIANT GLU-237.			
RX	MEDLINE=96414302; PubMed=8817330;			
RA	Hill M.R., Cookson W.O.;			
RT	"A new variant of the beta subunit of the high-affinity receptor for			
RL	immunoglobulin E (Fc epsilon RI-beta E237G): associations with			
RP	measures of atopy and bronchial hyper-responsiveness.";			
RL	Hum. Mol. Genet. 5:959-962(1996).			
RN	[4]			
RP	VARIANT GLU-237.			
RX	MEDLINE=96440420; PubMed=8842731;			
RA	Shirakawa T., Mao X.-Q., Sasaki S., Enomoto T., Kawai M., Morimoto K.,			
RT	Hopkin J.;			
RL	"Association between atopic asthma and a coding variant of Fc-epsilon-			
RP	RI-beta in a Japanese population.";			
RL	Hum. Mol. Genet. 5:1129-1130(1996).			
RN	[5]			
RP	ERRATUM.			
RX	MEDLINE=97123518; PubMed=8968765;			
RA	Shirakawa T., Mao X.-Q., Sasaki S., Enomoto T., Kawai M., Morimoto K.,			
RT	Hopkin J.;			
RL	Hum. Mol. Genet. 5:2068-2068(1996).			


```

RP SEQUENCE FROM N.A.
RX MEDLINE=86008235; PubMed=3900064;
RA Hoffman W.;
RT "Molecular characterization of the CAN1 locus in Saccharomyces cerevisiae. A transmembrane protein without N-terminal hydrophobic signal sequence.";
RL J. Biol. Chem. 260:11831-11837(1985).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=5288C / AB972;
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M., Chung E., Duncan A., Guzman E., Hartwell G., Hunnicke-Smith S., Hyman R., Kayser A., Komp C., Lastkati D., Lew H., Lin D., Moesdale D., Nakahara K., Namath A., Norgren R., Oelfner P., Oh C., Patel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V., Taylor P., Wei Y., Yelton M., Botstein D., Davis R.W.;
RA Submitted (DEC-1994) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: HIGH-AFFINITY PERMEASE FOR ARGINE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE AMINO ACID PERMEASE FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb.ch/announce/ or send an email to license@isb-sib.ch).
-----
MR EMBL: X03784; CAA27416.1; -
MR EMBL: M11724; AAA34467.1; -
MR EMBL: U18795; AAB65024.1; -
MR PIR: A23922; QRBYP.
MR SGD: S0000789; CAN1.
MR InterPro: IPR002027; -
DR Pfam: PF00324; aa_permeases; 1.
DR PROSITE: PS00218; AMINO_ACID_PERMEASE_1; 1.
KW Transport, Amino-acid transport, Transmembrane.
KW DOMAIN
FT TRANSMEM 93 110 POTENTIAL.
FT TRANSMEM 230 254 POTENTIAL.
FT TRANSMEM 324 343 POTENTIAL.
FT TRANSMEM 377 396 POTENTIAL.
FT TRANSMEM 420 437 POTENTIAL.
FT TRANSMEM 451 469 POTENTIAL.
FT TRANSMEM 494 518 POTENTIAL.
FT TRANSMEM 525 548 POTENTIAL.
FT TRANSMEM 549 590 POTENTIAL.
FT DOMAIN 534 534 HYDROPHILIC.
SEQUENCE 590 AA: 65785 MW: 455A21C77145330D CRC64;
-----
Query Match 10.6%; Score 81.5; DB 1; Length 590;
Best Local Similarity 22.9%; Pred. No. 2.5;
Matches 27; Conservative 26; Mismatches 42; Indels 23; Gaps 4;
OY 20 ASEVESTSLATFTSTQSPLOKLFARKKKILGTIOILFGIMTFSGVILFTLLKPY--P 77
Dy 294 AFTFGCTGLVGTIGCEANPRKSRVRAIK-----KVFPRILFTFYGLFTGLLWVNDP 348
OY 78 RFFPFLSGYPRWGVGLINSAGFLIAVKRKTTELGS-----LLITLMTSIIELFTI 129
Db 349 KLT-----QSTSYSTVSTPFIILAENSGTKVLPPIFNNAVILTLTIISANSNIYV 396
-----
RESULT 8
TCR_BACST STANDARD. PRT. 458 AA.
AC P07561.
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)

```

```

DE TETRACYCLINE RESISTANCE PROTEIN.
GN TET.
OS Bacillus stearothermophilus.
OG Plasmid pTH715.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
CC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxId=1442;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86031344; PubMed=2996983;
RA Hoshino T., Ikeda T., Tomizuka N., Furukawa K.;
RT "Nucleotide sequence of the tetracycline resistance gene of pTH715, a
RT thermophilic Bacillus plasmid: comparison with staphylococcal tcr
RT controls."
RL Gene 37:131-138(1985).
CC -I- FUNCTION: RESISTANCE TO TETRACYCLINE BY AN ACTIVE TETRACYCLINE
CC EFFLUX. THIS IS AN ENERGY-DEPENDENT PROCESS THAT DECREASES THE
CC ACCUMULATION OF THE ANTIBIOTIC IN WHOLE CELLS. THIS PROTEIN
CC FUNCTIONS AS A METAL-TETRACYCLINE/H+ ANTIPORTER.
CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -I- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
CC AS THE DRUG RESISTANCE TRANSLACASE FAMILY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, M1036; AAA22851.1; -.
DR PIR, A23973; YTBSTR.
DR InterPro, IPR001411; -.
DR PRINTS, PR01036; TCRTEB.
KW Antibiotic resistance; Transmembrane; Transport; Symport; Plasmid.
FT FT TRANSMM 12 33 POTENTIAL.
FT TRANSMM 81 100 POTENTIAL.
FT TRANSMM 111 129 POTENTIAL.
FT TRANSMM 140 162 POTENTIAL.
FT TRANSMM 165 185 POTENTIAL.
FT TRANSMM 201 221 POTENTIAL.
FT TRANSMM 223 240 POTENTIAL.
FT TRANSMM 256 276 POTENTIAL.
FT TRANSMM 297 317 POTENTIAL.
FT TRANSMM 324 344 POTENTIAL.
FT TRANSMM 346 365 POTENTIAL.
FT TRANSMM 432 451 POTENTIAL.
FT SEQUENCE 458 AA; 50119 MW; E1F334C57E7B06A CRC64;
SQ
Query Match 10.4%; Score 80; DB 1; Length 458;
Best Local Similarity 25.8%; Pred. No. 2.7;
Matches 32; Conservative 21; Mismatches 41; Indels 30; Gaps 6;
OY 32 TFSSTSPLOKLFARKMKILGTLQILFGIMTSFGYIFLTLLKYPKRPPIFLSGYPWG 91
Db 178 TITVTPFMKLLKRVKIRGHPDK-GITIMSVGIVF-----MPTTSYSI-- 223
OY 92 SVLFINSGAFILAVK--RKTP-----TLGILTLMTFSIIEFLSL-PSI 135
Db 224 SFLIVSVLSFLIFVAKIRKVDPFVDPGLGKNIPMIGVLCGILFGVYAGVSNVPM 283
OY 136 LGCH 139
Db 284 KDVH 287
RESULT 9
ID TCR_STRAG STANDARD; PRT; 458 AA.
NC P13924.
DT 01-JAN-1990 (Rel. 13, Created)

```

DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-OCT-1994 (Rel. 30, Last annotation update)
 DE TETRACYCLINE RESISTANCE PROTEIN.
 OS Streptococcus agalactiae.
 OG Plasmid pMV158.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 ON NCBI_TaxID=1311;
 RX MEDLINE=90016790; PubMed=2677995;
 RA van der Lelie D., Bron S., Venema G., Oskam L.;
 RT Similarity of minus origins of replication and flanking open reading
 frames of plasmids pUB110, pPB913 and pMV158.";
 RL Nucleic Acids Res. 17:7283-7294(1989).
 CC -1- FUNCTION: RESISTANCE TO TETRACYCLINE BY AN ACTIVE TETRACYCLINE
 EFFLUX. THIS IS AN ENERGY-DEPENDENT PROCESS THAT DECREASES THE
 ACCUMULATION OF THE ANTIBIOTIC IN WHOLE CELLS. THIS PROTEIN
 FUNCTIONS AS A METAL-TETRACYCLINE/H+ ANTI-PORTER.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
 AS THE DRUG RESISTANCE TRANSLOCASE FAMILY).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X15669; CAA3712.1;
 DR PIR: C25599; YTSOG.
 DR InterPro: IPR001411;
 DR PRINTS: PR01036; TCRTEB.
 KW Antibiotic resistance; Transmembrane; Transport; Symport; Plasmid.
 FT TRANSMEM 12 33 POTENTIAL.
 FT TRANSMEM 81 100 POTENTIAL.
 FT TRANSMEM 111 129 POTENTIAL.
 FT TRANSMEM 140 162 POTENTIAL.
 FT TRANSMEM 165 185 POTENTIAL.
 FT TRANSMEM 201 221 POTENTIAL.
 FT TRANSMEM 223 240 POTENTIAL.
 FT TRANSMEM 256 276 POTENTIAL.
 FT TRANSMEM 297 317 POTENTIAL.
 FT TRANSMEM 324 344 POTENTIAL.
 FT TRANSMEM 346 365 POTENTIAL.
 FT TRANSMEM 432 451 POTENTIAL.
 SO SEQUENCE 458 AA; 50006 MW; AD2014E7CA19995 CRC64;

Query Match 10.4%; Score 80; DB 1; Length 458;
 Best Local Similarity 25.8%; Pred. No. 2.7;
 Matches 32; Conservative 21; Mismatches 41; Indels 30; Gaps 6;

QY 32 TFSQSPQLQFAKMKILCTIQLFGIMTFSFGVIFLLKRPFRPFILSGYPFWG 91
 DB 178 TITVPLMKLKKVKIKGHEPIK-GILMSYGVF-----MLFTTSYGI-- 223
 QY 92 SVLFINGALILAVK--RKTE-----TIGILITLMTSEITLFIPL-PEST 135
 DB 224 SFLIVSVLSLIFVKHKKVTDPEVDPLGNILFMIGVLCGGIIFVAGFVSMVPYMM 283
 QY 136 LGCH 139
 DB 284 KDVA 287

RESULT 10
 TCR_STRPN
 ID TCR_STRPN STANDARD; PRT; 458 AA.
 AC P11063; P72219;

DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUL-1989 (Rel. 11, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE TETRACYCLINE RESISTANCE PROTEIN.
 OS Streptococcus pneumoniae; Bacillus cereus; and Bacillus subtilis.
 OG Plasmid pLS1, Plasmid pBC16, Plasmid pHY163PLK, Plasmid pPB19, and
 OG Plasmid pNS1981.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 ON NCBI_TaxID=1313, 1396, 1423;
 RX MEDLINE=90221899; PubMed=2109312;
 RA Paiva A., Vidgren G., Simonen M., Rintala H., Laamanen P.;
 RT "Nucleotide sequence of the tetracycline resistance gene of pBC16
 from Bacillus cereus.";
 RL Nucleic Acids Res. 18:1635-1635(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC PLASMID-PHY163PLK;
 RA Ishiwa H., Shibahara H.;
 RT "New shuttle vectors for Escherichia coli and Bacillus subtilis. III.
 Nucleotide sequence analysis of tetracycline resistance gene of
 pMn1p and ori-177.";
 RL Jpn. J. Genet. 60:485-496(1985).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC PLASMID-PHY163PLK;
 RA MEDLINE=83129391; PubMed=6186390;
 RT Selzer G., Som T., Itoh T., Tomizawa J.;
 RL "The origin of replication of plasmid pLSA and comparative studies on
 the nucleotide sequences around the origin of related plasmids.";
 RN Cell 32:119-129(1983).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC SPECIES-B. subtilis; PLASMID-pTB19;
 RA MEDLINE=92052681; PubMed=1946749;
 RT Oskam L., Hillenga D.J., Venema G., Bron S.;
 RL "The large Bacillus plasmid pTB19 contains two integrated
 rolling-circle plasmids carrying mobilization functions.";
 RN Plasmid 26:30-39(1991).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC SPECIES-B. subtilis; STRAIN-168 / MARBURG; PLASMID-PNS1981;
 RA MEDLINE=86287699; PubMed=3090576;
 RT Sakaguchi R., Shishido K., Hoshino T., Furukawa K.;
 RL "The nucleotide sequence of the tetracycline resistance gene of
 plasmid pNS1981 from Bacillus subtilis differs from pPHT15 from a
 thermophilic Bacillus by two base pairs.";
 RN Plasmid 16:72-73(1986).
 RN [7]
 RP -1- FUNCTION: RESISTANCE TO TETRACYCLINE BY AN ACTIVE TETRACYCLINE
 EFFLUX. THIS IS AN ENERGY-DEPENDENT PROCESS THAT DECREASES THE
 ACCUMULATION OF THE ANTIBIOTIC IN WHOLE CELLS. THIS PROTEIN
 FUNCTIONS AS A METAL-TETRACYCLINE/H+ ANTI-PORTER.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
 CC -1- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
 AS THE DRUG RESISTANCE TRANSLOCASE FAMILY).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial

```

RT      XIV from Saccharomyces cerevisiae between the BNI1 and the POL2
RL      genes."
RL      Yeast 12:505-514(1996).
CC      -1- FUNCTION: HIGH-AFFINITY PERMEASE FOR BASIC AMINO ACIDS.
CC      -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC      -1- SIMILARITY: BELONGS TO THE AMINO ACID PERMEASE FAMILY.
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
-----
DR      EMBL; X74069; CAA52199.1; -.
DR      EMBL; X92494; CAA63328.1; -.
DR      EMBL; Z71546; CAA96177.1; -.
DR      PIR; S44329; S44329.
DR      SGD; S0005214; ALP1.
DR      InterPro; IPR002027; -.
DR      Pfam; PF00324; aa-permeases; 1.
DR      PROSITE; PS00218; AMINO_ACID_PERMEASE_1; 1.
KW      Transport; Amino-acid transport; Transmembrane.
FT      TRANSMEM          75     95     POTENTIAL.
FT      TRANSMEM          96    116     POTENTIAL.
FT      TRANSMEM        154    174     POTENTIAL.

```

FT	TRANSMEM	266	286	POTENTIAL.
FT	TRANSMEM	307	327	POTENTIAL.
FT	TRANSMEM	360	380	POTENTIAL.
FT	TRANSMEM	404	424	POTENTIAL.
FT	TRANSMEM	432	452	POTENTIAL.
FT	TRANSMEM	478	498	POTENTIAL.
FT	TRANSMEM	511	531	POTENTIAL.
FT	CONFLICT	51		D -> DD (IN REF. 1).
FT	CONFLICT	126	126	V -> A (IN REF. 1).
FT	CONFLICT	260	260	D -> N (IN REF. 1).
FT	CONFLICT	517	517	I -> V (IN REF. 1).
FT	CONFLICT	548	548	R -> H (IN REF. 1).
SQ	SEQUENCE	573 AA;	64013 MW;	359DEP1466C348A1 CRC64;

OY	20	A	S	E	F	E	S	T	P	L	S	T	P	S	T	O	S	L	O	K	I	F	A	R	K	K	K	I	G	L	T	I	O	I	L	E	G	I	M	F	E	S	G	V	I	E	L	L	K	P	P	R	F	79
Db	277	A	E	T	T	G	G	E	L	V	C	T	T	G	E	A	N	P	K	A	L	P	R	A	K	---	K	V	A	R	I	L	V	E	T	I	L	S	E	F	T	G	L	V	P	-	N	D	330					
OY	80	P	E	I	F	L	S	G	P	E	M	S	V	L	E	I	N	S	G	A	F	L	I	A	R	K	R	T	E	L	G	---	I	L	T	M	F	S	I	E	L	F	I	129										
Db	331	P	K	L	S	D	S	G	---	I	F	V	S	S	P	E	M	S	I	E	N	S	G	K	V	L	P	D	I	F	N	A	V	L	L	I	L	S	A	G	S	N	V	I	379									
RESULT	12																																																					
ANSP_SALTY	ID	ANSP_SALTY	STANDARD;																																	PRT;	497	AA.																
AC	P40812;																																																					
DT	01-FEB-1995 (Rel. 31,	Created)																																																				
DT	01-FEB-1995 (Rel. 31,	Last sequence update)																																																				
DT	01-FEB-1995 (Rel. 31,	Last annotation update)																																																				
DE	L-ASPARAGINE PERMEASE (L-ASPARAGINE TRANSPORT PROTEIN).																																																					
GN	ANSP.																																																					
OS	Salmonella typhimurium.																																																					
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;																																																					
OX	NCBI_TaxID=602;																																																					
RN	[1]																																																					
RP	SEQUENCE FROM N.A.																																																					
RP	STRAIN=SA2656;																																																					

RX MEDLINE=95202072; PubMed=7894705;
 RA Jennings P.J., Anderson J.K., Beacham I.R.;
 RT "Cloning and molecular analysis of the *Salmonella enterica* ansp gene,
 RT encoding an L-asparagine permease.";
 RL Microbiology 141:141-146(1995).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE AMINO ACID PERMEASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U04851; AAA80001.1; -
 DR StyGene; SG10527; ansp.
 DR InterPro: IPR002027; -
 DR Pfam: PF00324; aa_permeases; 1.
 DR PROSITE: PS00218; AMINO ACID PERMEASE 1; 1.
 KM Transport: Amino-acid transport; Transmembrane; Inner membrane.
 FT TRANSMEM 34 54 POTENTIAL.
 FT TRANSMEM 58 78 POTENTIAL.
 FT TRANSMEM 109 129 POTENTIAL.
 FT TRANSMEM 146 166 POTENTIAL.
 FT TRANSMEM 171 191 POTENTIAL.
 FT TRANSMEM 219 239 POTENTIAL.
 FT TRANSMEM 264 284 POTENTIAL.
 FT TRANSMEM 298 318 POTENTIAL.
 FT TRANSMEM 353 373 POTENTIAL.
 FT TRANSMEM 378 398 POTENTIAL.
 FT TRANSMEM 422 442 POTENTIAL.
 FT TRANSMEM 448 468 POTENTIAL.
 SO SEQUENCE 497 AA; 54004 MW; BE80CB117992CBBB CRC64;
 Query Match 10.18; Score 77.5; DB 1; Length 497;
 Best Local Similarity 22.8%; Pred. No. 4.9;
 Matches 37; Conservative 25; Mismatches 47; Indels 53; Gaps 8;
 Oy 23 YESTELSATFTSTQSPLOKLFARKMKILGTQLFGIMTFSFGVIFLTLKPRPRF--- 79
 Db 233 FASIELVGTAGAECKDPKMKVPAKIN---SVIWRIGLFYGVSVLLVLLPNNAYOAG 287
 Oy 80 --PFI-FLS--GYPFMGSVLF-----NSG-----AFLLAVK 106
 Db 288 QSPVTFEFSKIGVPIYSIMIVLVLTALSLNSGLVCTGRIILKSMGSGAPFMAKMS 347
 Oy 107 RKTETUGILITLMTFSIIEFLIS-----LPPSLIG 137
 Db 348 ROHVPAAGIATLVVY-VGVFELNVLPVSRYEIVLNFASIG 388
 RESULT 13
 COTE_HUMAN STANDARD; PRT; 669 AA.
 ID COTE_HUMAN
 AC P81408;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE COTE1 PROTEIN.
 GN COTE1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Hipocampus;
 RA MEDLINE=97474796; PubMed=9331372;
 RT Winfield S.L., Tayebi N., Martin B.M., Gims E.I., Sidransky E.;
 RT "Identification of three additional genes contiguous to the

RT glucocerebrosidase locus on chromosome 1q21: implications for Gaucher
 RT disease.";
 RL Genome Res. 7:1020-1026(1997).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.
 CC -1- SIMILARITY: TO HUMAN KIA0574.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF023268; AAC51822.1; -
 DR Transmembrane.
 KM TRANSMEM 34 54 POTENTIAL.
 FT TRANSMEM 67 87 POTENTIAL.
 FT TRANSMEM 91 111 POTENTIAL.
 FT TRANSMEM 174 194 POTENTIAL.
 FT DOMAIN 244 250 POLY-PRO.
 FT DOMAIN 635 639 POLY-SER.
 SO SEQUENCE 669 AA; 71482 MW; E30360AC9A4571B6 CRC64;
 Query Match 10.08; Score 77; DB 1; Length 669;
 Best Local Similarity 22.6%; Pred. No. 7.2;
 Matches 31; Conservative 25; Mismatches 67; Indels 14; Gaps 5;
 Oy 16 PETTASEYESTELSATFTSTQSPLOKLFARKMKILGTQLFGIMTFSFGVIFLTLK 74
 Db 3 PPSDDSSRLTSRSTGTLRLHRPMLQALLGLVQLGLVYTFSMVASSVTTE 62
 Oy 75 PYPRFPFIIFLSGYPFMGSVLFINGAF-LIAVKKRTETGILITLMTFSIIEFLISLP 133
 Db 63 SIKR-----SCPSWAGFSLAFSGVGVSWKRPFLVISPFLS--SVLCVMISWAG 112
 Oy 134 SILGCHSEDC--DCEQC 148
 Db 113 SVLSCKNAQLARDPQC 129
 RESULT 14
 COTE_HUMAN STANDARD; PRT; 669 AA.
 ID COTE_HUMAN
 AC P29388;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 5 (EC 1.6.5.3).
 GN ND5 OR NAD5.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Mitochondrion.
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, COLUMBIA;
 RX MEDLINE=92007799; PubMed=1915303;
 RA Knop V., Schuster W., Wissinger B., Brennicke A.;
 RT "Trans splicing integrates an exon of 22 nucleotides into the nad5
 RT mRNA in higher plant mitochondria.";
 RL EMBO J. 10:3483-3493(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, COLUMBIA;
 RX MEDLINE=97141919; PubMed=8988169;
 RA Unseid M., Marientfeld J.R., Brandt P., Brennicke A.;
 RT "The mitochondrial genome of Arabidopsis thaliana contains 57 genes
 RT in 366,924 nucleotides.";

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2001, 04:07:35 ; Search time 40.14 Seconds
(without alignments)
300.552 Million cell updates/sec

Title: US-09-735-712-2

Perfect score: 1015
Sequence: 1 MDSTAHSPVFLVFPPEITA.....ISLPSTIGCHSDCCDCEQC 199

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

A.Geneseq_0601:*

- 1: /SIDSL/gcgdata/geneseq/geneseq/AA1980.DAT:*
- 2: /SIDSL/gcgdata/geneseq/geneseq/AA1981.DAT:*
- 3: /SIDSL/gcgdata/geneseq/geneseq/AA1982.DAT:*
- 4: /SIDSL/gcgdata/geneseq/geneseq/AA1983.DAT:*
- 5: /SIDSL/gcgdata/geneseq/geneseq/AA1984.DAT:*
- 6: /SIDSL/gcgdata/geneseq/geneseq/AA1985.DAT:*
- 7: /SIDSL/gcgdata/geneseq/geneseq/AA1986.DAT:*
- 8: /SIDSL/gcgdata/geneseq/geneseq/AA1987.DAT:*
- 9: /SIDSL/gcgdata/geneseq/geneseq/AA1988.DAT:*
- 10: /SIDSL/gcgdata/geneseq/geneseq/AA1989.DAT:*
- 11: /SIDSL/gcgdata/geneseq/geneseq/AA1990.DAT:*
- 12: /SIDSL/gcgdata/geneseq/geneseq/AA1991.DAT:*
- 13: /SIDSL/gcgdata/geneseq/geneseq/AA1992.DAT:*
- 14: /SIDSL/gcgdata/geneseq/geneseq/AA1993.DAT:*
- 15: /SIDSL/gcgdata/geneseq/geneseq/AA1994.DAT:*
- 16: /SIDSL/gcgdata/geneseq/geneseq/AA1995.DAT:*
- 17: /SIDSL/gcgdata/geneseq/geneseq/AA1996.DAT:*
- 18: /SIDSL/gcgdata/geneseq/geneseq/AA1997.DAT:*
- 19: /SIDSL/gcgdata/geneseq/geneseq/AA1998.DAT:*
- 20: /SIDSL/gcgdata/geneseq/geneseq/AA1999.DAT:*
- 21: /SIDSL/gcgdata/geneseq/geneseq/AA2000.DAT:*
- 22: /SIDSL/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	294	29.0	77	AA03074	Human secreted pro
2	206.5	20.3	239	AA096745	High affinity immu
3	206.5	20.3	239	AAV50174	Human high affinity
4	206.5	20.3	245	AAV94973	Human secreted pro
5	184.5	18.2	214	AAW06503	HTM4 protein. Hom
6	184.5	18.2	214	AAW41056	HTM4 protein. Hom
7	184.5	18.2	225	AAV94449	Human inflammation
8	184.5	18.2	242	AAV73495	Human secreted pro
9	184.5	18.2	248	AAV15225	Human receptor pro
10	184.5	18.2	248	AAV91531	Human secreted pro
11	184.5	18.2	248	AAW70489	Human hHA1ERDS-1so

12	184.5	18.2	250	AAV48505	Human breast tumor
13	184.5	18.2	257	AAW58419	Lung cancer associ
14	184.5	18.2	273	AAV91680	Human secreted pro
15	176.5	17.4	267	AAV15224	Human receptor pro
16	174.5	17.2	299	AAV91352	Human secreted pro
17	170	16.7	297	AAW20808	Human CD20 antigen
18	169	16.7	248	AAV6046	Extended human sec
19	168	16.6	297	AAV91356	CD20.4 antigen.
20	168	16.6	297	AAV96131	Human cell surface
21	167	16.5	297	AAW91436	Human CD20 antigen
22	167	16.5	297	AAW80445	Human CD20.4 anti
23	167	16.5	297	AAW86192	Human CD20.4 anti
24	162.5	16.0	237	AAW4585	Human HA1ERBS-iso
25	162.5	16.0	237	AAW70490	Human HA1ERBS-iso
26	162.5	16.0	240	AAV69998	Human HA1ERBS-iso
27	162	16.0	239	AAW4584	Human receptor-ass
28	162	16.0	249	AAW4583	Human HA1ERBS-iso
29	161.5	15.9	307	AAV91421	Human secreted pro
30	159	15.7	243	AAW05026	Beta subunit of ra
31	159	15.7	243	AAW42341	Beta subunit of the
32	159	15.7	246	AAW41770	Beta subunit of hl
33	159	15.7	246	AAW42337	Human FCER1 beta.
34	147.5	14.5	250	AAW61619	Clone HTPPE86 of T
35	147.5	14.5	250	AAW70432	Human cell surface
36	147.5	14.5	302	AAW54312	Human pancreatic c
37	142.5	14.0	167	AAV53632	A bone marrow secr
38	133	13.1	204	AAV91396	Human secreted pro
39	124.5	12.3	151	AAW54438	Mouse novel secret
40	124.5	12.3	151	AAW10245	Murine adult splee
41	123.5	12.2	244	AAW29149	Human high affinity
42	123.5	12.2	244	AAW75918	Human beta subunit
43	123.5	12.2	244	AAW74447	Human wild-type FC
44	123.5	12.2	244	AAW72900	Human IGERB SEQ ID
45	121.5	12.0	192	AAW74885	Human secreted pro

ALIGNMENTS

RESULT	1
AA03074	standard; Protein; 77 AA.
ID	AA03074:
AC	AA03074:
DT	06-OCT-2000 (first entry)
XX	Human secreted protein, SEQ ID NO: 7155.
DE	Human: 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX	gene therapy; chromosome mapping.
XX	Homo sapiens.
OS	EP1033401-A2.
PN	06-SEP-2000.
PD	21-FEB-2000; 2000EP-0200610.
XX	26-FEB-1999; 99US-0122487.
PR	(GEST) GENSET.
XX	Dumas Milne Edwards J, Duclert A, Giordano J;
PA	WPI: 2000-500381/45.
XX	N-PSDB; AAC03080.
DR	New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT	obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX	diagnostic, forensic, gene therapy and chromosome mapping procedures -

PS Claim 13; SEQ ID 7155; 71bp + CD-ROM; English.

XX The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30
 CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.

XX Sequence 77 AA;

SO Query Match 29.08; Score 294; DB 21; Length 77;
 Best Local Similarity 83.6%; Pred. No. 1.9e-27;
 Matches 56; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 33 ESTSPLOKLEFARKMKLTGTLQILFGIMTFSFGVLEFLTLKPPRPPIFLSGYPMGS 92
 Db 11 fglkapeckkllfarkmkllgtlqlfmxstfgvllfxltlkxprpiflksxptwgs 70

QY 93 VLFINSQ 99
 Db 71 vlfnsg 77

RESULT 2
 AAM96745
 ID AAM96745 standard; Protein; 239 AA.

XX AAM96745;
 AC 15-APR-1999 (first entry)
 DT 15-APR-1999 (first entry)
 XX High affinity immunoglobulin E receptor-like protein (IGERB).

DE High affinity immunoglobulin E receptor-like protein (IGERB).
 XX High affinity immunoglobulin E receptor-like protein; IGERB;
 KW Incyte clone 927955; inflammatory response; AIDS; Addison's disease;
 KW atherosclerosis; bronchitis; ulcerative colitis; diabetes mellitus;
 KW emphysema; gout; Graves's disease; osteoporosis; rheumatoid arthritis;
 KW Sjogren's syndrome; cancer.

XX Homo sapiens.
 XX Homo sapiens.

XX Key Location/Qualifiers
 FH Modified-site 125
 FT Modified-site /note= "potential protein kinase C phosphorylation
 FT site"
 FT Modified-site 137
 FT Modified-site /note= "potential N-glycosylation site"
 FT Modified-site 192
 FT Modified-site /note= "potential casein kinase II phosphorylation
 FT site"

XX US5871930-A.
 PN 16-FEB-1999.
 PD 16-FEB-1999.
 XX 21-AUG-1997; 97US-0916902.
 PF 21-AUG-1997; 97US-0916902.
 XX 21-AUG-1997; 97US-0916902.
 PR 21-AUG-1997; 97US-0916902.
 XX (INCY-) INCYTE PHARM INC.
 PA (INCY-) INCYTE PHARM INC.
 XX Bandman O, Corley NC, Lal P;
 PI Bandman O, Corley NC, Lal P;
 XX WPI; 1999-166628/14.
 DR

DR N-PSDB; AAX15104.
 XX New polynucleotide coding for high affinity immunoglobulin E
 PT receptor-like protein - and its complement, useful for diagnosis,
 PT prevention and antisense therapy of inflammatory responses

XX Claim 1; Fig 1A-D; 30pp; English.

PS The present sequence represents a high affinity immunoglobulin E
 XX receptor-like protein (IGERB). The present sequence was first
 CC identified in Incyte clone 927955 for the brain cDNA library
 CC BRAIN0704. Polynucleotides complementary to the IGERB cDNA can
 CC be used as probes to IGERB gene expression. IGERB proteins and
 CC nucleotides may be used for the diagnosis of inflammatory
 CC responses associated with expression of IGERB, e.g. AIDS,
 CC Addison's disease, atherosclerosis, bronchitis, ulcerative colitis,
 CC diabetes mellitus, emphysema, gout, Graves's disease, osteoporosis,
 CC rheumatoid arthritis, Sjogren's syndrome and complications of cancer.
 CC The sequences may also be useful in assays that detect activation or
 CC induction of various cancers.

XX Sequence 239 AA;

SO Query Match 20.3%; Score 206.5; DB 20; Length 239;
 Best Local Similarity 31.8%; Pred. No. 2.1e-16;
 Matches 49; Conservative 32; Mismatches 70; Indels 3; Gaps 2;

QY 39 LQKLFAR-KMKILGTQILFGIMTFSFGVLEFLTLKPPRPPIFLSGYPMGSVLEFIN 97
 Db 54 lqekflkgepkvlgvdlalmslmgllmmcnasntysnpslygylvsgvnmfl 113

QY 98 SGAFILAVKRRKTEETLILSRIMNLALRAIAGIILTF--GFLDONTICGYSHONSQ 155
 Db 114 sgslstaaglrctkglyrsglsmitsvslaagllntslafsthrhpcyngysnsm 173

QY 156 CKAATVLEFLGILTLMTFSITELFISLPFSILGC 189
 Db 174 chgtmslmgldgmvlilsvlefclavslaafgc 207

RESULT 3
 AAY50174
 ID AAY50174 standard; Protein; 239 AA.

XX AAY50174;
 AC 31-JAN-2000 (first entry)
 DT 31-JAN-2000 (first entry)
 XX Human high affinity IGE receptor-like protein (IGERB).

DE Human high affinity IGE receptor-like protein (IGERB).
 XX IGERB; high affinity immunoglobulin E; IGE; receptor; allergy;
 KW inflammation; B lymphocyte; homology; high affinity IGE receptor; FCRI;
 KW mast cell; basophil; histamine; protease; cytoplasmic granule; synthesis;
 KW effector; prostaglandin; leukotriene; cytokine; antigen; antibody;
 KW identification; agonist; antagonists; expression; activity; diagnosis;
 KW therapy; inflammatory disorder; multiple sclerosis; osteoarthritis;
 KW asthma; cancer; side effect; complication.

XX Homo sapiens.
 XX Homo sapiens.

XX Key Location/Qualifiers
 FH Modified-site 125
 FT Modified-site /note= "phosphorylated by protein kinase C"
 FT Modified-site 137
 FT Modified-site /note= "glycosylated"
 FT Modified-site 192
 FT Modified-site /note= "phosphorylated by casein kinase II"

XX US5977072-A.
 PN 02-NOV-1999.
 PD 02-NOV-1999.
 XX

OY SCAPLIAVARKTETTEIIRSRINLLSLRAALGILLTF--GEILDQNTICGRSHONSQ 155
||:|||::| | ||: |: :|| ::| |
Db 120 sgslsaagirttkglvrslygmicsvsylaasgillntislafsfyfhpcnygnsmn 179

OY 156 CKAATVLFLGLITLTMFTSFIEELFISLPSSILGC 189
|::|:|::|::|::|:
Db 180 chgtmsllngldgmwillsvlefciafvlsatfsgc 213

RESULT 5
AAW06503 standard; protein; 214 AA.
XN AAW06503;
XX AC AAW06503;
XT DT 06-FEB-1997 (first entry)
● HTm4 protein.

KM Human HTm4 protein; Fc(epsilon)RI receptor; FR; atopic disease;
KW allergy; asthma; atopic dermatitis; allergic rhinitis; hereditary.
XX Homo sapiens.
OS Key Location/Qualifiers
FX Domain /note= "Transmembrane domain"
FT Domain /note= "Transmembrane domain"
FT Domain /note= "Transmembrane domain"
FT Domain /note= "Transmembrane domain"
FT Domain /note= "Transmembrane domain"

PX US5552312-A.
PN 03-SEP-1996.
PD 06-OCT-1994; 94US-0318492.
PX 06-OCT-1994; 94US-0318492.
PR 06-OCT-1994; 94US-0318492.
XX (BETH-) BETH ISRAEL HOSPITAL BOSTON.
PA Adra CN, Leijas J, Lim B;

PL WPI; 1996-412066/41.
DR N-PDSDB; AAT45120.
XX New isolated mammalian HT(M4) protein DNA - used to develop prods.
PT for use in the diagnosis and treatment of atopic diseases, e.g.
PT allergy or asthma
TX Claim 2; Fig 1; 10pp; English.

CC This sequence represents human HTm4 protein. HTm4 is related to the
CX Fc(epsilon)RI receptor (FR). HTm4 contains four hydrophobic domains
XX or 20-21 amino acids. The N-terminal region before the first
CC hydrophobic domain contains four prolins residues, and each of the
CC hydrophilic regions between the hydrophobic regions contains one
CC prolins. This protein crosses the plasma membrane four times
CC forming two small extracellular loops and has both the N- or C-
CC terminals in the cytoplasm. The HTm4 gene and protein can be used
CC in the research and study of the induction of expression of FR and
CC the function of its subunits. They can be used in the design of
CC drugs which can block or inhibit induction of FR, thereby treating atopical
CC diseases such as allergy, asthma, atopic dermatitis and allergic rhinits.
CC The HTm4 gene sequences may also be used to detect patients suffering
CC from hereditary atopic disorders and to map genes on the human
CC chromosome.

XZ Sequence 214 AA:

Query Match	Score	18.2%	Score 184.5	DB 17	Length 214
Best Local Similarity	28.2%		Pred. No. 7.3e-14		
Matches 59	Conservative 39	Mismatches 86	Indels 25	Gaps	8
Oy	1	MDSSTAH-SPVELVPEPPTTASEYESTELSATFTSTOSPLOKLFARKKKILCTIQLFGI	59		
Db	11	lgasahagpgpsetgpeelntsvh-----plngspdyk---aklqvlgaqlnaa	60		
Oy	60	MTFSGVGFILFLLPKPY----PRPFPIPLSGYPPMGSVLFINSGAFLAVKKKTETLII	115		
Db	61	mla1gav-flgslgypynhgnkhnffllytygyp1gwavficsgslsvaagikprtwlq	119		
Oy	116	LSRIKMLLSALRAIGIILLTFPGFIIDQNYI--CGYSHONSQ----CKAVTVLFGLILT	169		
Db	120	nsfgmiasatalaygtalislinaivnigslsc---hsseepalcnyngs1ngmws1	176		
Oy	170	IMTFSIIEFLFISLPFSILGCHSEDCDCQ	198		
Db	177	lllllelcvltistiamwcnancnre	205		
RESULT	6				
AAW41056					
ID	AAW41056	standard; Protein: 214 AA.			
XX	AAW41056;				
AC					
XX					
DT	29-APR-1998	(first entry)			
XX					
DE	HTM4 protein.				
XX					
HM	HTM4; antibody; Ige Fc receptor; fceps1lonribeta; CD20 antigen; TRAF-1;				
KM	4-transmembrane spanning protein superfamily; ligand binding mimic;				
KW	haematopoietic cell detection; inhibitor; tumour necrosis factor; KAP;				
KM	TNF receptor-associated factor; CDK-activating kinase; TRAF-2; TRAF-3;				
KW	TM4SF.				
XX					
OS	Homo sapiens.				
XX					
PN	US5705615-A.				
XX					
PD	06-JAN-1998.				
XX					
PF	03-SEP-1996;	96US-0707340.			
XX					
PR	03-SEP-1996;	96US-0707340.			
XX					
PR	06-OCT-1994;	94US-0318492.			
XX					
PR	03-JUL-1996;	96US-0675648.			
XX					
PA	(BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.				
XX					
XX	Adra CN, Lelias J, Llm B;				
PI					
DR	WPI. 1998-086194/08.				
XX					
PI	N-PSDB; AAV03875.				
XX					
PT	Antibody specific for HTM4 protein - useful to block or mimic ligand				
XX	binding or detect haematopoietic cells				
PS	Disclosure: Column 25-28; 18pp; English.				
XX					
CC	This sequence is the HTM4 protein. The HTM4 protein is specifically				
CC	bound by the antibody of the invention. HTM4 shows homology to the beta				
CC	subunit of the high affinity Ige Fc receptor (fceptslonribeta) and the				
CC	CD20 antigen, both of which are in the 4-transmembrane spanning proteins				
CC	superfamily (TM4SF). The antibody of the invention is used to block or				
CC	mimic binding of ligands such as the tumour necrosis factor (TNF)				
CC	receptor-associated factors TRAF-1, TRAF-2 and TRAF-3 and the				
CC	phosphatase, CDK-activating kinase (KAP), to receptors comprising HTM4,				
CC	and to detect haematopoietic cells.				
XX					

PT Novel polynucleotides and proteins having biological activities which
 PT make them suitable for treating, preventing or ameliorating medical
 PT conditions in humans or animals -
 XX
 XX
 PS Claim 221: Page 718; 730pp; English.
 XX
 CC The present invention describes human secreted proteins encoded by
 CC polynucleotides obtained from adult testes, foetal brain, adult brain,
 CC brain (foetal and adult), foetal kidney, adult spleen, and adult thymus
 CC cDNA libraries. The polynucleotides and proteins are predicted to have
 CC biological activities which would make them suitable for treating,
 CC preventing or ameliorating medical conditions in humans and animals.
 CC Suggested activities include nutritional activity, cytokine and cell
 CC proliferation/differentiation activity, immune stimulating (e.g. as
 CC vaccines) or suppressing activity, haematopoiesis regulating activity,
 CC tissue growth activity, activin/inhibin activity, chemotactic/
 CC chemokinetic activity, hemostatic and thrombolytic activity, receptor/
 CC ligand activity, anti-inflammatory activity, cadherin/tumour invasion
 CC suppressor activity, and tumour inhibition activity. The polynucleotides
 CC are also stated to be useful for gene therapy. Therapeutic compositions
 CC are also presently valuable for veterinary applications. AA252475 to
 CC AA252581 encode human secreted proteins, and AAY73390 to AAY73500
 CC represent human secreted proteins, given in the present invention.
 CC
 XX
 SQ Sequence 242 AA;
 Query Match 18.2%; Score 184.5; DB 21; Length 242;
 Best Local Similarity 26.4%; Pred. No. 8.9e-14;
 Matches 55; Conservative 40; Mismatches 90; Indels 23; Gaps 6;
 QY 1 MDSSTAHSPFLVPPPEITASEVESTELSATFTSTQSPLOKLFARKKKILGTIQLFGIM 60
 Db 1 mtsqpyneitvipsnvl--nfsqaekpeptngqdsikkhlaekivglitqlcgmm 58
 QY 61 TFSRGVFLFTLKP-YRPPFIRL-SGYPPGWSVLINSAGFLIAVKRKTTELLILSR 118
 Db 59 vlsigillaasfsptitqvtstlinsayptigpffifisqslstakekrtikilvns1 118
 QY 119 IMNLLSLRAIVAGIIL-----TFGFIIDPN-----YICGSHON---SCKAV 159.
 Db 119 vgsilalsalsavglilsvkqatlnpaslqceloknhiprsvsyfthdslytdcyta 178
 QY 160 TVFLGLITLMTFSITELFSLPFSIL 187
 Db 179 kaslagclsimlctlliefclavltavl 206
 RESULT 9
 AAY15225
 ID AAY15225 standard; protein: 248 AA.
 XX
 XX AAY15225:
 AC
 XX 26-OCT-1999 (first entry)
 DT
 XX
 DE Human receptor protein (HURP) 4 amino acid sequence.
 XX
 XX receptor; cancer; autoimmune disorder; inflammation;
 KM antagonist; cell surface protein; cell signalling;
 KM antibody; human receptor protein; HURP; reproductive disorder;
 KM developmental disorder; gastrointestinal disorder.
 XX
 XX Homo sapiens.
 OS
 XX
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 8 /note- "Potential N-glycosylation site"
 FT Misc-difference 20 /note- "Potential N-glycosylation site"
 FT Misc-difference 74 /note- "Potential N-glycosylation site"
 FT Misc-difference 110 /note- "Potential N-glycosylation site"

FT /note- "Potential CAMP-/cGMP-dependent protein-
 FT kinase phosphorylation site"
 FT Misc-difference 22 /note- "Potential casein kinase II-
 FT phosphorylation site"
 FT Misc-difference 193 /note- "Potential casein kinase II-
 FT phosphorylation site"
 FT Misc-difference 36 /note- "Potential protein kinase C-
 FT phosphorylation site"
 FT Misc-difference 105 /note- "Potential protein kinase C-
 FT phosphorylation site"
 FT Misc-difference 136 /note- "Potential protein kinase C-
 FT phosphorylation site"
 FT Misc-difference 177 /note- "Potential protein kinase C-
 FT phosphorylation site"
 FT Misc-difference 233 /note- "Potential protein kinase C-
 FT phosphorylation site"
 FT Misc-differences 242 /note- "Potential tyrosine kinase
 FT phosphorylation"
 XX
 PN W09941375-A2.
 XX
 PD 19-AUG-1999.
 XX
 PF 05-FEB-1999; 99WO-US02572.
 XX
 PR 12-FEB-1998; 98US-0022939.
 XX
 PA (INCY-) INCYTE PHARM INC.
 XX
 PI Au-Young J, Bandman O, Baughn M, Corley NC, Guegler KJ;
 PI Hillman JL, Lal P, Shah P, Tang YT, Yae H;
 XX WPI: 1999-494536/41.
 DR N-PSDB; AA206369.
 DR
 XX
 XX New human receptor proteins, used e.g. to treat, prevent and
 PT diagnose gastrointestinal and developmental disorders - and related
 PT nucleic acids, vectors, transformed cells, antibodies, agonists and
 PT antagonists
 PS Claim 1; Page 81-82; 94pp; English.
 XX
 XX The human receptor protein 4 (HURP-4) has 22% homology with the
 CC rat IGE receptor and 19% homology with human CD20 protein.
 CC HURP-4 is expressed in cancerous, inflamed, hematopoietic/immune
 CC and gastrointestinal tissue. HURP-4 therefore appears to have a role in
 CC some forms of cancer, autoimmune/inflammatory disorders, and
 CC gastrointestinal disorders.
 CC This gives rise to the possibility of using an antagonist or an antibody
 CC of HURP to treat or prevent cancer or autoimmune/inflammatory disorders.
 CC
 XX
 SQ Sequence 248 AA;
 Query Match 18.2%; Score 184.5; DB 20; Length 248;
 Best Local Similarity 26.4%; Pred. No. 8.9e-14;
 Matches 55; Conservative 40; Mismatches 90; Indels 23; Gaps 6;
 QY 1 MDSSTAHSPFLVPPPEITASEVESTELSATFTSTQSPLOKLFARKKKILGTIQLFGIM 60
 Db 1 mtsqpyneitvipsnvl--nfsqaekpeptngqdsikkhlaekivglitqlcgmm 58
 QY 61 TFSRGVFLFTLKP-YRPPFIRL-SGYPPGWSVLINSAGFLIAVKRKTTELLILSR 118
 Db 59 vlsigillaasfsptitqvtstlinsayptigpffifisqslstakekrtikilvns1 118

	CC	osteoporosis, arthritis, infections, AIDS, spinal cord injuries,
	CC	transplant rejection, diabetes, asthma, sepsis, acne, psoriasis,
	CC	cardiovascular disorders, reproductive disorders, gastrointestinal
	CC	disorders, respiratory disorders and metabolic disorders. The proteins
	CC	or polymucleotides can also be used as food additives or preservatives.
	CC	The proteins are also useful for identifying their binding partners.
	CC	AA626337 to AA626345 and AA91450 are sequences used in the
	CC	exemplification of the present invention.
	XX	
	XX	Sequence 248 AA;
	XX	
	Query Match	18.2%; Score 184.5; DB 21; Length 248;
	Best Local Similarity	26.4%; Pred. NO. 8.9e-14;
	Matches 55; Conservative	40; Mismatches 90; Indels 23; Gaps 6;
OY	1	MDSSTVAHPVLPVPPEPTATASEYEESTETLSAATFTFSQSPLOKLFARKMKLTIGTITGFCIM 60
Db	1	mtsgvqpettiivpsnvi--nfsgaekepeplngqdsdkkhkhaelkvigtldlgcmm 58
OY	61	TTFSEGVFLFTLLKP-YRFPPIFL-SCYPWGSVLFNTSGAFLIAVRKRTTEPLIISR 118
Db	59	vlsjgilliasasfsfpfcqvctllnsaypfifffilissglstakekrilkilvhse1 118
OY	119	IMNIIISARAIAGIRLL-----TFGFILQN-----YICGYSHON---SCKAV 159
Db	119	vgsilisalsalvgfiliilsvkqatlnpaslgeidknnpitrsyvsylyhdsltytocyta 178
OY	160	TVEFGILTTLMTFESIIEFLISLPPSIL 187
Db	179	kaslagnxismliclliecfclavltavl 206
	RESULT 11	
	AAB70489	
ID	AAB70489 standard; Protein:	248 AA.
AC	AAB70489;	
XX		
DT	04-MAY-2001 (first entry)	
XX		
DE	Human hHAIERBs-Iso protein sequence SEQ ID NO:7.	
KW	Human: hHAIERBs-Iso; HAIERBs; HAIERBs isomer; detection;	
OS	high affinity immunoglobulin epsilon receptor beta subunit.	
XX	Homo sapiens.	
XX	CNI269410-A.	
XX	11-OCT-2000.	
XX	17-MAR-2000; 2000CN-0114959.	
XX	17-MAR-2000; 2000CN-0114959.	
XX	(SREN-) SOUTHERN RES CENT NAT HUMAN GENE GROUP.	
XX	Xiao H, Liu F, Song H;	
XX	WPI: 2001-050545/07.	
DR	N-PSDB; AAF63724.	
PT	New human immunoglobulin receptor subunit protein and its nucleic acid	
XX		
PS	Claim 4; Page 17; 22pp; Chinese.	
XX		
CC	The present invention describes a human high affinity immunoglobulin	
CC	epsilon receptor beta subunit isomer, designated hHAIERBs-Iso.	
CC	hHAIERBs-Iso is isolated from in human pheochromocytoma. The present	
CC	invention also describes methods for the preparation and detection of	
CC	hHAIERBs-Iso protein and nucleotide sequences. The present sequence	

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 7, 2001, 04:07:40 ; Search time 24.66 Seconds
(without alignments)
166.159 Million cell updates/sec

Title: US-09-735-712-2

Perfect score: 1015

Sequence: 1 MDSTAHSPVFLVFPPEITA.....ISLPSTLGHSEDCDCQC 199

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCYTUS_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/Backfilesl.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	206.5	20.3	239	2	US-08-916-902A-1 Sequence 1, Appl1
2	206.5	20.3	239	2	US-09-213-389-1 Sequence 1, Appl1
3	184.5	18.2	214	1	US-08-318-492-4 Sequence 4, Appl1
4	184.5	18.2	214	2	US-08-707-340-4 Sequence 4, Appl1
5	184.5	18.2	214	2	US-08-916-902A-3 Sequence 3, Appl1
6	184.5	18.2	214	2	US-08-994-578-4 Sequence 4, Appl1
7	184.5	18.2	214	2	US-09-213-389-3 Sequence 3, Appl1
8	159	15.7	243	1	US-07-869-933-29 Sequence 29, Appl1
9	159	15.7	243	1	US-08-201-879A-4 Sequence 4, Appl1
10	159	15.7	243	2	US-08-916-902A-4 Sequence 4, Appl1
11	159	15.7	243	2	US-08-916-902A-4 Sequence 4, Appl1
12	159	15.7	243	2	US-09-213-389-4 Sequence 4, Appl1
13	159	15.7	243	4	US-09-103-663-29 Sequence 29, Appl1
14	159	15.7	243	4	US-09-103-663-33 Sequence 33, Appl1
15	159	15.7	246	4	US-07-869-933-23 Sequence 23, Appl1
16	159	15.7	246	4	US-09-103-663-23 Sequence 23, Appl1
17	158.5	15.6	235	1	US-07-869-933-34 Sequence 34, Appl1
18	158.5	15.6	235	1	US-08-201-879A-5 Sequence 5, Appl1
19	158.5	15.6	235	4	US-09-103-663-34 Sequence 34, Appl1
20	123.5	12.2	244	1	US-07-869-933-32 Sequence 32, Appl1
21	123.5	12.2	244	4	US-08-201-879A-3 Sequence 3, Appl1
22	123.5	12.2	244	4	US-09-103-663-32 Sequence 32, Appl1
23	85	7.6	327	4	US-08-748-506-24 Sequence 24, Appl1
24	80.5	7.9	311	3	US-08-605-284B-23 Sequence 23, Appl1
25	80.5	7.9	1956	4	US-08-843-417-2 Sequence 1, Appl1
26	79.5	7.8	241	3	US-08-808-148-1 Patent No. 5386025
27	79	7.8	1872	6	5386025-6

28	79	7.8	1873	1	US-08-435-675B-4 Sequence 4, Appl1
29	78.5	7.7	423	2	US-08-494-907-14 Sequence 14, Appl1
30	78.5	7.7	423	5	PCT-US96-10986-14 Sequence 14, Appl1
31	78.5	7.7	956	2	US-08-897-443-3 Sequence 3, Appl1
32	78	7.7	347	4	US-09-097-889-14 Sequence 14, Appl1
33	77.5	7.6	1956	4	US-08-843-417-10 Sequence 10, Appl1
34	77	7.6	1416	1	US-08-061-465-4 Sequence 4, Appl1
35	76	7.5	1873	1	US-08-336-257A-7 Sequence 7, Appl1
36	75	7.4	459	4	US-08-097-889-22 Sequence 22, Appl1
37	73.5	7.2	381	1	US-08-467-125-2 Sequence 2, Appl1
38	73.5	7.2	381	2	US-08-911-320A-2 Sequence 2, Appl1
39	73.5	7.2	381	4	US-09-217-101-2 Sequence 4, Appl1
40	73.5	7.2	398	1	US-08-097-938-4 Sequence 4, Appl1
41	73.5	7.2	398	1	US-08-476-000-4 Sequence 4, Appl1
42	73.5	7.2	398	1	US-08-472-840-4 Sequence 4, Appl1
43	73.5	7.2	398	2	US-08-476-976-4 Sequence 4, Appl1
44	73.5	7.2	398	3	US-08-474-410-4 Sequence 4, Appl1
45	73	7.2	371	2	US-08-928-692-20 Sequence 20, Appl1

ALIGNMENTS

RESULT 1
US-08-916-902A-1
; Sequence 1, Application US/08916902A
; Patent No. 5871930
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
; TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; Zip: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/916,902A
; FILING DATE: Herewith
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0371 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 239 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT04
; CLONE: 927955
; US-08-916-902A-1

RESULT 4

US-08-707-340-4

; Sequence 4, Application US/08707340

; Patent No. 5705615

; GENERAL INFORMATION:

; APPLICANT: Lim, Bing

; APPLICANT: Adra, Chaker N.

; APPLICANT: Lejias, Jean-Michel

; TITLE OF INVENTION: RECOMBINANT HTM4 GENE, PROTEIN AND

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

; STREET: Two Militia Drive

; CITY: Lexington

; STATE: Massachusetts

; COUNTRY: U.S.A.

; ZIP: 02173

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/707,340

; FILING DATE: 03-SEP-1996

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/318,492

; FILING DATE: 06-OCT-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/675,648

; FILING DATE: 03-JUL-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Brook, David E.

; REGISTRATION NUMBER: 22,592

; REFERENCE/DOCKET NUMBER: B1H94-03A2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 861-6240

; TELEFAX: (617) 861-9540

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 214 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-707-340-4

Query Match 18.2%; Score 184.5; DB 1; Length 214;

Best Local Similarity 28.2%; Pred. No. 3.3e-13;

Matches 59; Conservative 39; Mismatches 86; Indels 25; Gaps 8;

QY 1 MDSSTAH-SVPLVFPPEITASEYESTELSATPTSTQSPLOKLFARKMKLTIGTIIQIFGI 59

DB 11 LGSASAHGTGSGETGPELMTSVYH-----PINGSPTYQK--AKLYVLAIGAIQIIMAA 60

QY 60 MTFSEGVIFLFTLLKPY----PREPFIISGYFPWGSVLFINSQAFLLIAVKRKTETFLII 115

DB 61 MILALGV-FIGSIQYRPHQKHEFFFTYGYPIWGAVFCCSGTSLSVAGIKPTRMIQ 119

QY 116 LSRMNLISALRAIAGIILITFGFIIDONTI--CGYSHQSQ-----CKATVLEFLGIT 169

DB 120 NSFGMNIASATIALVGAFLSLNIAVNIQSLRSC---HSSSESPDLGNYGSIISNGVSL 176

QY 170 LMTFSIIEFLISPLFSILGSHEDCDEQ 198

DB 177 LLITLLELCVTTISTIAMMCNANCNSRE 205

RESULT 5

; Sequence 4, Application US/08994578

; Patent No. 5871930

; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga

; APPLICANT: Lal, Preeti

; APPLICANT: Corley, Neil C.

; TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incycle Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; OPERATING SYSTEM: IBM compatible

; SOFTWARE: FASTSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/916,902A

; FILING DATE: Herewith

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0371 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

; TELEFAX: 415-845-4166

; TELEX:

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 214 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: GenBank

; CLONE: 561639

US-08-916-902A-3

Query Match 18.2%; Score 184.5; DB 2; Length 214;

Best Local Similarity 28.2%; Pred. No. 3.3e-13;

Matches 59; Conservative 39; Mismatches 86; Indels 25; Gaps 8;

QY 1 MDSSTAH-SVPLVFPPEITASEYESTELSATPTSTQSPLOKLFARKMKLTIGTIIQIFGI 59

DB 11 LGSASAHGTGSGETGPELMTSVYH-----PINGSPTYQK--AKLYVLAIGAIQIIMAA 60

QY 60 MTFSEGVIFLFTLLKPY----PREPFIISGYFPWGSVLFINSQAFLLIAVKRKTETFLII 115

DB 61 MILALGV-FIGSIQYRPHQKHEFFFTYGYPIWGAVFCCSGTSLSVAGIKPTRMIQ 119

QY 116 LSRMNLISALRAIAGIILITFGFIIDONTI--CGYSHQSQ-----CKATVLEFLGIT 169

DB 120 NSFGMNIASATIALVGAFLSLNIAVNIQSLRSC---HSSSESPDLGNYGSIISNGVSL 176

QY 170 LMTFSIIEFLISPLFSILGSHEDCDEQ 198

DB 177 LLITLLELCVTTISTIAMMCNANCNSRE 205

RESULT 6

; Sequence 4, Application US/08994578

Patent No. 5972688
GENERAL INFORMATION:
APPLICANT: Lim, Bing
APPLICANT: Adra, Chaker N.
APPLICANT: Leillas, Jean Michel
TITLE OF INVENTION: HTM4 METHODS OF TREATMENT AND ASSAYS,
TITLE OF INVENTION: AGONISTS AND ANTAGONISTS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/994,578
FILING DATE: December 19, 1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/707,340
FILING DATE: 03-SEP-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/675,648
FILING DATE: 03-JUL-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/318,492
FILING DATE: 06-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: BIH94-03A22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 861-6240
TELEFAX: (781) 861-9540
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 214 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-994-578-4

Query Match 18.2%; Score 184.5; DB 2; Length 214;
Best Local Similarity 28.2%; Pred. No. 3.3e-13;
Matches 59; Conservative 39; Mismatches 86; Indels 25; Gaps 8;

QY 1 MDSSTAH-SVPLVFPPEITASESESTELSATTFSTQSPLOKLFARKMKILGTQILFGI 59
DB 11 LGSASAHGTGSGETGPELNTSVH-----PINGSPTYOK--AKLOYGAIQIILNAA 60
QY 60 MTSFVGIVFFTLTKPY-----PREPFIILSGYPFGWSVLFINSGAFILAVKRRKTTETLLI 115
DB 61 MILALGV-FLGSLOYRPHQKHFFFTFYGYPIWGAVFCCSGTSLSVAGIKPTRWIO 119
QY 116 LSRIMNLSALRAIAGIILLTFGFIIDONYI--CGYSHONSQ---CKAVTVLFLGILIT 169
DB 120 NSFCOMINASATIALVGAFLSLINAVNIQSLRSC---HSSSESPDLCLNYGSIISNGVSL 176
QY 170 LMTFSIIEFLISLPSILGCHSEDCOEQ 198
DB 177 LLITLLELCVTISTIAMCMNANCNSRE 205

RESULT 7
US-09-213-389-3
Sequence 3, Application US/09213389

Patent No. 5977072
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/213,389
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/916,902
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0371 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 214 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: 561639
US-09-213-389-3

Query Match 18.2%; Score 184.5; DB 2; Length 214;
Best Local Similarity 28.2%; Pred. No. 3.3e-13;
Matches 59; Conservative 39; Mismatches 86; Indels 25; Gaps 8;

QY 1 MDSSTAH-SVPLVFPPEITASESESTELSATTFSTQSPLOKLFARKMKILGTQILFGI 59
DB 11 LGSASAHGTGSGETGPELNTSVH-----PINGSPTYOK--AKLOYGAIQIILNAA 60
QY 60 MTSFVGIVFFTLTKPY-----PREPFIILSGYPFGWSVLFINSGAFILAVKRRKTTETLLI 115
DB 61 MILALGV-FLGSLOYRPHQKHFFFTFYGYPIWGAVFCCSGTSLSVAGIKPTRWIO 119
QY 116 LSRIMNLSALRAIAGIILLTFGFIIDONYI--CGYSHONSQ---CKAVTVLFLGILIT 169
DB 120 NSFCOMINASATIALVGAFLSLINAVNIQSLRSC---HSSSESPDLCLNYGSIISNGVSL 176
QY 170 LMTFSIIEFLISLPSILGCHSEDCOEQ 198
DB 177 LLITLLELCVTISTIAMCMNANCNSRE 205

RESULT 8
US-07-869-933-29
Sequence 29, Application US/07869933
Patent No. 5770396
GENERAL INFORMATION:

type: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-201-879A-4

Query Match	15.7%	Score 159	DB 1	Length 243
Best Local Similarity	28.0%	Pred. No. 2.9e+10		
Matches 45	Conservative 29	Mismatches 71	Indels 16	Gaps 3

```

0Y 36 QSPLOKLEARKKMLTGTQLQLEFGMTFSFGVILETLKp--YRPFIFILSGYPMGVS 93
Db 47 QOTWOSPELKKLEFLGAYQYVLVGLICCFGVWCSTQTSDEDEVLLLYRAGYPMFGAV 106
0Y 94 LEINSGAFLVLRKRTETTLILSRINMLLSALRAINGILLTFEGFLDQNYIOGYSHON 155
Db 107 LEVLISGELISMERKNTLYLRGSLGANIVYSIAAGLGIALILLINSSNAYM----- 159
0Y 154 SOCKAVT-----VLFLGLIITLMPFSIIEFLPSLPSII 187
Db 160 MYCKDITEDDCGFVTSFTTELVLALLPELITLAFSAAVLLII 200

```

RESULT 11
 US-08-916-902A-4
 ; Sequence 4, Application US/08916902A
 ; Patent No. 5871930
 ; GENERAL INFORMATION:
 APPLICANT: Bandman, Olga
 APPLICANT: Lal, Preeti
 TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
 TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Drive
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/916,902A
 FILING DATE: Herewith
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.

```

1 REGISTRATION NUMBER: 36,749
2 REFERENCE/DOCKET NUMBER: PF-0371 USA
3
4 TELECOMMUNICATION INFORMATION:
5
6 TELEPHONE: 415-855-0555
7
8 TELEFAX: 415-845-4166
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
10
```

Query Match	15.7%	Score 159	DB 2	length 243
Best Local Similarity	28.0%	Pred. No. 2.9e-10		
Matches 45	Conservative 29	Mismatches 71	Indels 16	Gaps 3

```

0Y 36 QSPLOKLEARKMKILGTIOLIEGIMTFSFGVILFTLLKP--YRPPFIETLSGYPFGW 93
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 47 QQIMWSEFLKKELEFLGATQVLAWGLICCFGTVCSTQSDSEDEVLLLEYRAGIPFGAV 106
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
0Y 94 LEINSGAEFLIIVKRRKTEETLILSRIMNLALRAIAGIILTPGFIIDQNYIGYSHQN 153
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 107 LEVLSGFLSIWKRRNTIYIVRGSGLAIVYSIAAGIGIALIILINISNNSAYM----- 159
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
0Y 154 SCKAVT-----VFLGILITLTMTFSIIEFLTSLPFSIL 187
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 160 NYCKDITEDGCEVYTSFTTELVALELLEPTLIAPGSAVLLII 200
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT . 12
 US-09-213-389-4
 Sequence 4, Application US/09213389
 Patent No. 5977072
 GENERAL INFORMATION:
 APPLICANT: Bandman, Olga
 APPLICANT: Lal, Preeti
 APPLICANT: Corley, Neil C.
 TITLE OF INVENTION: HIGH AFFINITY IMMUNOGLOBULIN E
 TITLE OF INVENTION: RECEPTOR-LIKE PROTEIN
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.,
 STREET: 3174 Porter Drive
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: PASTSEQ for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/213,389
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/916,902
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0371 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 TELEX:

```

1 NAME: BENT, Stephen A.
2 REGISTRATION NUMBER: 29,768
3 REFERENCE/DOCKET INFORMATION: 40399/154 NIND
4 TELECOMMUNICATION INFORMATION:
5 TELEPHONE: (703)836-9300
6 TELEFAX: (703)683-4109
7 TELEX: 899149
8 INFORMATION FOR SEQ ID NO: 23:
9 SEQUENCE CHARACTERISTICS:
10 LENGTH: 246 amino acids

```

TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-869-933-23

Query Match 15.78; Score 159; DB 1; Length 246;
Best Local Similarity 28.0%; Pred. No. 2.9e-10;
Matches 45; Conservative 29; Mismatches 71; Indels 16; Gaps 3;

QY 36 QSPLOKLFARRMKITGFIQILFGIMTFSPGVIFLFLKLP--YRPPPIFLSGYPFWGSV 93
Db 50 QQTWQSFLEKLELELVGVTVQVAVGLICFEGIVVCSLTQTSDEDEVLLLYRAGYPEWGAV 109
QY 94 LFINSAGFLIAVKRKTEETLILSRIMNLSALRAIAGIILITFGFIIDQNYICGYSHON 153
Db 110 LFVLSGFLSIMSERKNTLYVRSGISGANIVYSIAAGIAIILILNLSNNSAYM----- 162
154 SOCKAVT-----VFLGLITIMTFSITELFISLPSPIL 187
163 NYCKDITEDDGCFTVSFITEIVMLFLFTLIAFCSAVILLII 203

Search completed: October 7, 2001, 04:40:00
Job time: 1940 sec

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	466.2	77.3	508	16	AI149899 gfa3h06.x
2	456.6	74.7	516	7	AA436088 zuo3a08.r
3	347	57.5	387	7	AA416972 zt94h05.s
4	340.2	56.4	382	7	AA435988 zuo3a08.s
5	291.8	48.4	394	1	AA758635 ab67b04.s
6	283.8	47.1	484	163	BE107659 UT-R-BT1
7	267.4	44.3	410	7	AA411806 zt67a03.s
8	245.4	40.7	389	11	AA781801 at58g01.s
9	242.8	40.3	415	7	AA470059 zt94h05.r
10	207	34.3	482	14	AI002083 ct38b02.s
11	141	23.4	332	147	BF319786 uy63h11.x
12	131.8	21.9	601	138	BE638317 EST00003
13	126.2	20.9	224	138	BE638325 EST00002
14	115.6	19.2	538	224	AO108532 CTF-HSP-2
15	113	18.7	281	11	AA707529 ah41a12.s
16	86.4	14.3	1071	106	AL544561 AL544561
17	84	13.9	793	154	BG484817 602505511
18	83	13.8	218	138	BE638321 EST00013
19	82.2	13.6	704	155	BG571626 602502934
20	80.4	13.3	508	4	AA234138 zt51b06.r
21	80.4	13.3	562	7	AA418443 zt92e05.x
22	70.6	11.7	421	257	B86842 RPC111-26F1
23	61.2	10.1	887	106	AL531049 AL531049
24	59.4	9.9	715	155	BG538851 602568296
25	55.6	9.2	579	136	BE513276 601315340
26	55.6	9.2	947	132	AK008652 Mus muscu
27	55.4	9.2	470	102	AI807884 w43g11.x
28	55	9.1	528	188	BE675149 7103402.x
29	55	9.1	172	602	BG024663 602275459
30	54.6	9.1	697	143	BF056859 7K10g10.x
31	54.4	9.0	1504	132	AK003110 Mus muscu
32	53	8.8	1184	132	AK017928 Mus muscu
33	52.6	8.7	515	153	BG434337 602506331
34	52.6	8.7	594	32	AV717594 AV717594
35	52.6	8.7	597	32	AV716310 AV716310
36	52.6	8.7	619	32	AV661743 AV661743
37	52.6	8.7	680	155	BG546745 602574115
38	52.6	8.7	735	32	AV715678 AV715678
39	52.4	8.7	504	150	BF522968 UT-R-C2P-
40	52.4	8.7	755	134	AI950360 wp10b06.x
41	52.2	8.7	736	137	BE563951 601348160
42	52.2	8.7	860	106	AL551751 AL551751
43	51.4	8.5	534	151	BF591089 7h52g02.x
44	51.4	8.5	559	19	AI391038 mc10h04.y
45	50.8	8.4	482	151	BF600732 265501 MA

ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE
1	AI149899/c	508 bp mRNA	EST	10-NOV-1998						
	gfa3h06.x1	Soares-testis-NHT Homo sapiens cDNA clone IMAGE:1752827								
	3'	similar to SW-CD20_HUMAN P11836 B-LYMPHOCTE ANTIGEN CD20 ;								
		mRNA sequence.								
	AI149899									
	AI149899.1	GI:3678368								
	EST									
	human.									
	Homo sapiens									
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.									
	1 (bases 1 to 508)									
	NCI-CGAP htcp://www.ncbi.nlm.nih.gov/ncicgap.									
	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),									
	Tumor Gene Index									

JOURNAL COMMENT

Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps-rt@mail.nih.gov
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/dbp/image/image.html
 Insert Length: 742 Std Error: 0.00
 Seq primer: -40m13 fwd. EF from Amersham
 High quality sequence stop: 455.
 Location/Qualifiers
 1..508
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1752827"
 /clone_1ib="Soares-testis-NHT"
 /sex="male"
 /lab_host="DH10B"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGGCCGCCCAATTTTCTTTTCTTTT 3']
 Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73D vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT

186 a 91 c 89 g 142 t

ORIGIN

Seq ID NO: 1 alignment with AI149899
 Query Match 77.3% Score 466.2; DB 16; Length 508;
 Best Local Similarity 99.4%; Pred. No. 5e-109; Mismatches 3; Indels 0; Gaps 0;
 Matches 468; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY	133	agaaaaatgaatcttagagcatccagatcctgttggaattatgaccttctt	192
DB	508	agaaaaatgaatcttagagcatccagatcctgttggaattatgaccttctt	449
QY	193	gaagatctctcttccatcttgtaaaacacatccaaaggttcccttattctt	252
DB	448	gaagatctctcttccatcttgtaaaacacatccaaaggttcccttattctt	389
QY	253	taagatatccatcttggtggtctgtttgttcattatctcggaaccttcaattg	312
DB	388	taagatatccatcttggtggtctgtttgttcattatctcggaaccttcaattg	329
QY	313	gtgaaagaagaaacacagaaactcgtataattagccgaataatgaactcttact	372
DB	328	gtgaaagaagaaacacagaaactcgtataattagccgaataatgaactcttact	269
QY	373	gccctagagcaatagctgaatcattccctcacatcttggttccatccatgaac	432
DB	268	gccctagagcaatagctgaatcattccctcacatcttggttccatccatgaac	209
QY	433	tacatttggttattctccacaaatagtcagtgtaagctgttactgtctgtctg	492
DB	208	tacatttggttattctccacaaatagtcagtgtaagctgttactgtctgtctg	149
QY	493	ggaatttgatcatttggttacttcacatattatgaatttttctctctgccttc	552
DB	148	ggaatttgatcatttggttacttcacatattatgaatttttctctctgccttc	89
QY	553	tcatttggtggtccaccacagagattgttgtaacaaatgtgttga 603	
DB	88	tcatttggtggtccaccacagagattgttgtaacaaatgtgttga 38	

THIS PAGE BLANK (USPTO)

AA336088 516 bp mRNA EST 09-NOV-1997
 Z00308.01 Soares_testis_NRT Homo sapiens cDNA clone IMAGE:730742
 5' similar to SW:CD20_HUMAN P11836 B-LYMPHOCYTE ANTIGEN CD20 ;
 mRNA sequence.

ACCESSION: AA336088
 VERSION: AA336088.1 GI:2141002
 EST.

KEYWORDS: human.

SOURCE: Homo sapiens
 ORGANISM: Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 516)
 Hillier, L., Allen, M., Bowles, L., Dubuque, T., Gelsel, G., Jost, S.,
 Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin,
 J., Moore, B., Schellenberg, K., Stepien, M., Tan, F., Thelising, B.,
 White, Y., Wyllie, T., Waterston, R. and Wilson, R.
 WashU-NCI human EST Project
 Unpublished (1997)
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

CCOUNT
 TITLE
 JOURNAL

FEATURES
 source
 1..516
 /organism="Homo sapiens"
 /db_xref="GDB:592743"
 /db_xref="taxon:9606"
 /clone="IMAGE:730742"
 /clone_11b="Soares_testis_NRT"
 /sex="male"
 /lab_host="DH10B"
 /note="vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from Clontech Laboratories,
 Inc., and primed with a Not I - oligo(dT) primer [5'
 TGTTCACATCTGAGTGGAGCGCGCCGACCAATTTTCTTTTCTTTT 3']
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT7T3 vector. Library
 went through one round of normalization to Cot5, and was
 constructed by Bento Soares and M. Fatima Bonaudo."

BEST
 COUNT
 ORIGIN
 149 a 117 c 80 g 170 t

Query Match 74.7%; Score 450.6; DB 7; Length 516;
 Best Local Similarity 99.1%; Pred. NO. 4.9e-105;
 Matches 453; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

0Y 1 atgggttaagacgcgcacacagtcgcggtgtttcgtgtattccctccgaataactgct 60
 Db 52 ATGGGATTCAAGACCGCACACAGTCGCGGTGGTATTTTCTTCCCAAAATCACTGCT 111

0Y 61 tcaagatatgagtcacagaaacttcacgcacagaccttcaactcaagaagcccttgcaa 120
 Db 112 TCAGGATATGAGTCACAGAACTTTCAGCGCAGAGACTTTCAACTCAAGCCCTTGCAA 171

0Y 121 aaataattgctagaaaaatgaaaactcttaggagctatccagatccctgttggaaatg 180
 Db 172 AAATATTTGCTAGAAAATGAAAATCTTAGGAGACTATCCAGATCCTGTTGGAATTATG 231

0Y 181 accttctcttggtggttatctctcttccacctgttaaaacccatccaaggtttccc 240
 Db 232 ACCTTCTCTTGTGGAGTATCTCTCTTTCACTTTGTGTTAAACCATATCCAAAGGTTTCCC 291

Query Match	Best Local Similarity	Matches	350; Conservative
57.58;	Score 347;	DB 7;	Length 387;
98.68;	Pred. No. 1.3e-71;		
0;	Mismatches 5;		Indels 0; Gaps 0;

FEATURES	SOURCE
1. 387	Location/Qualifiers
/organism="Homo sapiens"	
/db_xref="GDB:5926570"	
/db_xref="taxon:9606"	
/clone="IMAGE:730041"	
/clone_id="Soares_testis_NHT"	
/sex="male"	
/lab_host="DH10B"	
/note="Vector: pUT73D-Pac (Pharmacia) with a modified polylinker; Site:1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCCCAATTTTATTTTATTTT 3'] Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pUT73 vector. Library went through one round of normalization to Cots, and was constructed by Benito Soares and M. Fatima Bonaldo."	
138 a	77 c 67 g 105 t

FEATURES	SOURCE
1. 387	Location/Qualifiers
/organism="Homo sapiens"	
/db_xref="GDB:5926570"	
/db_xref="taxon:9606"	
/clone="IMAGE:730041"	
/clone_id="Soares_testis_NHT"	
/sex="male"	
/lab_host="DH10B"	
/note="Vector: pUT73D-Pac (Pharmacia) with a modified polylinker; Site:1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCCCAATTTTATTTTATTTT 3'] Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pUT73 vector. Library went through one round of normalization to Cots, and was constructed by Benito Soares and M. Fatima Bonaldo."	
138 a	77 c 67 g 105 t

FEATURES	SOURCE
1. 387	Location/Qualifiers
/organism="Homo sapiens"	
/db_xref="GDB:5926570"	
/db_xref="taxon:9606"	
/clone="IMAGE:730041"	
/clone_id="Soares_testis_NHT"	
/sex="male"	
/lab_host="DH10B"	
/note="Vector: pUT73D-Pac (Pharmacia) with a modified polylinker; Site:1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCCCAATTTTATTTTATTTT 3'] Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pUT73 vector. Library went through one round of normalization to Cots, and was constructed by Benito Soares and M. Fatima Bonaldo."	
138 a	77 c 67 g 105 t

FEATURES	SOURCE
1. 387	Location/Qualifiers
/organism="Homo sapiens"	
/db_xref="GDB:5926570"	
/db_xref="taxon:9606"	
/clone="IMAGE:730041"	
/clone_id="Soares_testis_NHT"	
/sex="male"	
/lab_host="DH10B"	
/note="Vector: pUT73D-Pac (Pharmacia) with a modified polylinker; Site:1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCCCAATTTTATTTTATTTT 3'] Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pUT73 vector. Library went through one round of normalization to Cots, and was constructed by Benito Soares and M. Fatima Bonaldo."	
138 a	77 c 67 g 105 t

FEATURES	SOURCE
1. 387	Location/Qualifiers
/organism="Homo sapiens"	
/db_xref="GDB:5926570"	
/db_xref="taxon:9606"	
/clone="IMAGE:730041"	
/clone_id="Soares_testis_NHT"	
/sex="male"	
/lab_host="DH10B"	
/note="Vector: pUT73D-Pac (Pharmacia) with a modified polylinker; Site:1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCCCAATTTTATTTTATTTT 3'] Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pUT73 vector. Library went through one round of normalization to Cots, and was constructed by Benito Soares and M. Fatima Bonaldo."	
138 a	77 c 67 g 105 t

FEATURES	SOURCE
1. 387	Location/Qualifiers
/organism="Homo sapiens"	
/db_xref="GDB:5926570"	
/db_xref="taxon:9606"	
/clone="IMAGE:730041"	
/clone_id="Soares_testis_NHT"	
/sex="male"	
/lab_host="DH10B"	
/note="Vector: pUT73D-Pac (Pharmacia) with a modified polylinker; Site:1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCCCAATTTTATTTTATTTT 3'] Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pUT73 vector. Library went through one round of normalization to Cots, and was constructed by Benito Soares and M. Fatima Bonaldo."	
138 a	77 c 67 g 105 t

FEATURES	SOURCE
1. 387	Location/Qualifiers
/organism="Homo sapiens"	
/db_xref="GDB:5926570"	
/db_xref="taxon:9606"	
/clone="IMAGE:730041"	
/clone_id="Soares_testis_NHT"	
/sex="male"	
/lab_host="DH10B"	
/note="Vector: pUT73D-Pac (Pharmacia) with a modified polylinker; Site:1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCCCAATTTTATTTTATTTT 3'] Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pUT73 vector. Library went through one round of normalization to Cots, and was constructed by Benito Soares and M. Fatima Bonaldo."	
138 a	77 c 67 g 105 t

FEATURES	SOURCE
1. 387	Location/Qualifiers
/organism="Homo sapiens"	
/db_xref="GDB:5926570"	
/db_xref="taxon:9606"	
/clone="IMAGE:730041"	
/clone_id="Soares_testis_NHT"	
/sex="male"	
/lab_host="DH10B"	
/note="Vector: pUT73D-Pac (Pharmacia) with a modified polylinker; Site:1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCCCAATTTTATTTTATTTT 3'] Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pUT73 vector. Library went through one round of normalization to Cots, and was constructed by Benito Soares and M. Fatima Bonaldo."	
138 a	77 c 67 g 105 t

THIS PAGE BLANK (USPTO)